Optimisation Study of Geographical Allocation of Solar PV Generation Capacity for Grid Support

by Vicki-Mari Vermeulen

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr. JM Strauss

Department of Electrical and Electronic Engineering

December 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2016

Abstract

Due to growing concerns regarding environmental impacts and energy security, the renewable energy sector and its role in large-scale power generation is becoming increasingly significant, both locally and internationally. While it reduces reliance on fossil fuels, however, grid-integration of renewable energy generation also introduces considerable risks at high penetration levels due to the intermittent nature of sources such as photovoltaic (PV) solar and wind energy. The development of methodologies for addressing these concerns has become a prominent field of research, which includes aspects such as smart-grid technologies and optimal siting and sizing of renewable energy plants with regard to the grid infrastructure and demand profile. This project investigates one such aspect, namely optimisation of the allocation of solar PV generation capacity in the context of grid-related metrics such as the seasonal and daily peak demand periods and energy supply variability. The main research objectives focus on determining the impacts of optimal geographical distribution of solar PV sources in the South African context, using grid support metrics defined in terms of diurnal and seasonal trends as performance criteria.

The development and implementation of the optimisation strategy required the identification and specification of appropriate optimisation parameters, such as sets of potential locations that reflect the seasonal and diurnal diversity of the local solar resource, a comprehensive range of objective functions that is representative potential grid support metrics and a range of optimisation algorithms. Three groups of locations, referred to as site groups, were selected based on geographical significance with regard to diurnal and seasonal solar cycles. A set of optimisation objectives were defined based on grid-support considerations such as maximising average energy delivery, prioritising delivery during peak demand periods and minimising the day-to-day variability of the solar PV renewable energy contribution. A selection of candidate metaheuristic optimisation algorithms were identified for evaluation, namely the genetic algorithm, two standard variations of the pattern search technique and an additional pattern search variation that incorporates a genetic algorithm for a hybrid approach. The various site groups, diurnal and seasonal specifications, optimisation objectives and optimisation algorithms were consolidated to define a comprehensive set of optimisation problem cases.

The evaluation and subsequent analysis of the problem cases was implemented via an integrated software platform that forms part of an ongoing software application development project. Based on the designated optimisation strategy, a solar PV optimisation module was developed and integrated with an established database-driven user interface. The correspond-

ing relational database structure was optimised for all current and future applications. The simulation software for implementing the optimisation problem cases on an external simulation platform was developed and amalgamated with the solar PV optimisation module.

The results of the optimisation study confirm the seasonal and diurnal significance of the geographical distribution of solar PV generation, with seasonal variation occurring along a north-south axis and diurnal variation occurring along an east-west axis. The results for the same optimisation objectives evaluated for different seasonal and diurnal periods often show disparities. This indicates that solar PV distributions that exhibit the best performance characteristics overall are not necessarily ideal for supporting peak demand periods. This supports the notion that variable feed-in tariffs rather than the commonly used flat tariffs should apply to renewable energy generation, since this would encourage development that supports the grid without sacrificing plant profitability. The results for the different optimisation objectives also show a clear trade-off between maximising the annual cumulative energy yield and minimising variability of supply. In general, the daily variability throughout a seasonal period decreases as geographical dispersion increases, often at the cost of lowering the cumulative annual energy yield. With regard to the optimisation algorithms investigated, the technique combining pattern search and the genetic algorithm proved to be the most robust. Due to its non-deterministic nature, however, the algorithm generally necessitated multiple evaluations to ensure high quality solutions.

The quantitative impacts of the results achieved in the study are, to a degree, limited due to the geographical parameters of the South African case study considered in the investigation. The proposed optimisation strategy, however, has excellent potential in the context of larger interconnected grids, such as the European grid, United States mainland and the Southern African power pool region where an increased range of diurnal and seasonal characteristics applies. With regard to future work, it is recommended that a similar optimisation strategy should be investigated for combined wind and solar PV generation, since the contrasting characteristics of these sources could produce much more optimal aggregated power profiles.

Opsomming

As gevolg van groeiende kommer rakende omgewingsimpak en energie-sekuriteit word die hernubare energiesektor en die rol daarvan in grootskaalse kragopwekking tans van toenemend belang, plaaslik sowel as internasionaal. Alhoewel dit afhanklikheid van fossielbrandstowwe verminder, bied die netwerkintegrasie van hernubare energie-opwekking by hoë penetrasievlakke ook aansienlike risikos as gevolg van die onderbroke aard van bronne soos fotovoltaïese (PV) sonkrag en windkrag. Die ontwikkeling van metodes vir die hantering van hierdie probleme het 'n prominente navorsingsgebied geword, wat aspekte soos slimnetwerk tegnologie en optimale plasing en vermoëns van hernubare energieaanlegte met betrekking tot die kragstelsel infrastruktuur en aanvraagprofiel insluit. Hierdie projek ondersoek een van hierdie aspekte, naamlik optimisering van die toekenning van PV opwekkingskapasiteit in die konteks van netwerkverwante eienskappe soos die seisoenale en daaglikse spitsaanvraag periodes en variasie in energieverskaffing. Die belangrikste navorsingsdoelwitte fokus op bepaling van die impakte van optimale geografiese verspreiding van PV sonkragbronne in die Suid-Afrikaanse konteks, deur gebruik te maak van netwerkondersteunende kriteria soos gedefinieer word in die konteks van daaglikse en seisoenale tendense as prestasiekriteria.

Die ontwikkeling en implementering van die optimiseringstrategie het die identifisering en spesifikasie van gepaste optimiseringsparameters, soos stelle van potensiële liggings wat die seisoenale en daaglikse diversiteit van die sonkrag hulpbron reflekteer, 'n omvattende reeks van optimeringsfunksies wat verteenwoordigend is van netwerk ondersteunings kriteria en reeks van optimiseringsalgoritmes. Drie groepe van liggings, waarna verwys word as liggingsgroepe, is gekies op grond van geografiese belang met verwysing na daaglikse en seisoenale sonsiklusse. 'n Stel van optimeringsdoelwitte is gedefinieer, gebaseer op netwerk ondersteuningsoorwegings soos maksimering van gemiddelde energielewering, prioritisering van lewering gedurende spitstyd aanvraagperiodes en die mimimering van die dag-tot-dag wisselvalligheid van die PV hernubare energiebydrae. 'n Seleksie van metaheuristiese optimeringsalgoritmes is geïdentifiseer vir evaluering, naamlik die genetiese algoritme, twee standaard variasies van die patroonsoek tegniek en 'n bykomende patroonsoek variasie wat 'n genetiese algoritme vir 'n hibriede benadering inkorporeer. Die verskillende liggingsgroepe, daaglikse en seisoenale spesifikasies, optimeringsdoelwitte en optimeringsalgoritmes is gekonsolideer om 'n omvattende stel optimerings probleemgevalle te definieer.

Die evaluering en daaropvolgende ontleding van die probleemgevalle is geïmplementeer deur middel van 'n geïntegreerde sagtewareplatform wat deel vorm van 'n deurlopende sagteware

ontwikkelingsprojek. Gebaseer op die aangewese optimiseringstrategie, is ’n PV sonkrag optimiseringsmodule ontwikkel en met die gevestigde databasis-gedrewe gebruikerskoppelvlak geïntegreer. Die relasionele databasisstruktuur is geoptimiseer vir alle huidige en toekomstige toepassings. Die simulasiesagteware vir die evaluering van die optimerings probleemgevalle is op ’n eksterne simulasieplatform ontwikkel en met die PV son optimiseringsmodule gekombineer.

Die resultate van die optimiseringstudie bevestig die seisoenale en daaglikse belangrikheid van die geografiese verspreiding van PV sonkragopwekking, met seisoenale variasie wat voorkom langs ’n noord-suid-as en daaglikse variasie wat voorkom langs ’n oos-wes-as. Die resultate vir dieselfde optimiseringsdoelwitte, geëvalueer vir verskillende seisoenale en daaglikse periodes, toon dikwels verskille. Dit dui daarop dat PV sonverspreidings wat die beste algehele prestasie eienskappe toon nie noodwendig ideaal is vir die ondersteuning van spitsaanvraag periodes nie. Dit ondersteun die idee dat veranderlike invoertariewe eerder as die algemeen gebruikte plat tariewe van toepassing moet wees vir hernubare energieopwekking, aangesien dit ontwikkeling sal aanmoedig wat die netwerk ondersteun sonder om die winsgewendheid van aanlegte te benadeel. Die resultate vir die verskillende optimiseringsdoelwitte toon ook ’n duidelike uitruil tussen maksimering van die kumulatiewe jaarlikse energieopbrengs en mimimering van variansie in voorsiening. In die algemeen neem die daaglikse variansie deur die loop van ’n seisoenale tydperk af soos die geografiese verspreiding verhoog, dikwels ten koste van n verlaging in die kumulatiewe jaarlikse energieopbrengs. Met betrekking tot die optimiseringsalgoritmes wat ondersoek is, is bevind dat die die gekombineerde patroonsoek en genetiese algoritme tegniek die mees robuuste is. As gevolg van die nie-deterministiese aard daarvan het die algoritme verskeie evaluerings vereis om kwaliteit oplossings te verseker.

Die kwantitatiewe impakte van die resultate behaal in die studie is, tot ’n mate, beperk as gevolg van die geografiese parameters van die Suid-Afrikaanse gevallestudie wat in die ondersoek aangespreek is. Die voorgestelde optimeringsstrategie het egter uitstekende potensiaal in die konteks van groter tussenverbinde netwerke, soos die Europese netwerk, Verenigde State vasteland en die Suider-Afrikaanse kragpoel netwerk waar n wyer reeks van daaglikse en seisoenale karakteristieke van toepassing is. Met betrekking tot verdere werk word aanbeveel dat ’n soortgelyke optimiseringstrategie ondersoek word vir gesamentlike wind en PV sonkrag opwekking, aangesien die kontrasterende eienskappe van hierdie bronne baie meer optimale kumulatiewe kragprofiele kan produseer.

Acknowledgements

I would like to express my sincere gratitude towards the following parties for their invaluable support – whether academical, financial or otherwise – throughout the duration of this project:

- My supervisor, Dr. JM Strauss
- The National Research Foundation
- My parents
- My friends, especially George Blignault and the other alumni of Batteljon E222

Contents

List of Tables

List of Figures

Acronyms

1NP	1st Normal Form
2NP	2nd Normal Form
3NP	3rd Normal Form
CSP	concentrated solar power
CSV	comma-separated values
CV	coefficient of variation
DBMS	database management system
DE	Differential Evolution
DG	distributed generation
DLL	dynamic-link library
DNI	Direct Normal Irradiance
DoE	Department of Energy
ERD	entity relationship diagram
FK	foreign key
GA	genetic algorithm
GDP	gross domestic product
GHG	greenhouse gas
GHI	Global Horizontal Irradiance
GPS	generalised pattern search
GSS	generating set search
GTI	Global Tilted Irradiance
GUI	graphical user interface

IDE	integrated development environment
IPP	independent power producer
IRP	Integrated Resource Plan
kWh	kilowatt-hour
NDP	National Development Plan
PK	primary key
PV	photovoltaic
RDM	relational database model
RE	renewable energy
REFITs	renewable energy feed-in tariffs
REIPPPP	Renewable Energy Independent Power Producer Procurement Program
RSD	relative standard deviation
RVC	rapid voltage change
SD	standard deviation
SDK	software development kit
SQL	Structured Query Language
TMY	typical meteorological year
TOU	Time-of-Use
VCL	Visual Component Library

Chapter 1

Introduction

1.1 Renewable energy in South Africa

In recent years the global energy market has encountered critical long-term challenges such as environmental pollution, climate change and diminishing fossil fuel resources. This has given rise to extensive growth in renewable energy generation capacity and technologies throughout the world due to increased emphasis on sustainable development and energy security through diversification of generation capacity [1]. South Africa is no exception to this trend and has seen the development of a significant local renewable energy sector during the past five years, which can largely be attributed to changes in energy policy due to environmental concerns coupled with the urgent need for increased generation capacity [2].

Since the initial implementation of local utility-scale power generation, the South African power system has consisted predominantly of large coal-fired power stations situated close to the mines and industries of the inland provinces of Gauteng, Mpumalanga and Limpopo, with generation and distribution almost entirely owned and controlled by the publicly owned national power utility Eskom [3]. Since 2004, however, Eskom's power reserve margins have declined rapidly due to the growth of the national electricity demand overtaking the substantial excess of generation capacity installed over the previous 20 years. This growing threat to energy security necessitated the implementation of mitigating actions on both the demand and supply side, leading to significant increases in electricity tariffs as well as investment in the construction of two massive new coal-fired power stations [2]. By 2008, constraints on the power system had become so severe that Eskom investigated load shedding to preserve the stability of the national grid. Despite these mitigation measures, however, plant availability has been decreasing since 2009, mainly as a consequence of breakdowns and other maintenance-related problems [4]. The increase in unplanned outages is attributed to factors such as the following:

- Long-term deterioration of maintenance quality.
- Delaying critical maintenance on plants in order to meet daily energy demand as reserves declined.

- Extended restoration times on base-load capacity plants, since the majority of installations are past their mid-life.

- Additional maintenance requirements due to the impacts of declining coal quality on plant performance.

- Disruption of fuel supply to power stations.

In addition to these occurrences, the implementation of the new coal-fired plants suffered severe delays, which further increased the risks to grid stability and diminished the reliability of Eskom's power generation capacity. Since a strong correlation exists between a reliable and adequate electricity supply and positive economic growth [5], it is not surprising that the South African economy also suffered significant adverse effects under these conditions. The need for energy security to support economic development in a country already fraught with socio-economic challenges consequently served as a strong motivator for policy makers to push for accelerated implementation of local renewable generation capacity.

With regard to environmental impacts, the strong reliance on coal in South Africa's mining-based, energy-intensive economy results in an unusually high level of CO_2 emissions per capita and relative to the gross domestic product (GDP) [2]. Consequently, even though South Africa was not bound by any obligations to reduce greenhouse gas (GHG) emissions under the 1992 Kyoto protocol or the United Nations Framework Convention on Climate Change, the Department of Environmental Affairs initiated research on long-term mitigation strategies. This initiative eventually led to South Africa pledging CO_2 emissions reductions of 34 % by 2020 and 42 % by 2025 (subject to proper support from the international community) at the 2009 Copenhagen Conference of Parties [6]. At the 2011 COP17 meeting in Durban, public and private sector stakeholders further consented to a list of commitments designed to realise the government's aim of creating 300 000 green economy jobs by 2020.

The aforementioned pledges led to the introduction of carbon emissions limitations as well as the inclusion of renewable energy (RE) targets in the most recent Integrated Resource Plan (IRP) (issued for the period 2010-2030 and last updated in 2013) [7], even though these steps increased the overall cost of the energy plan [6]. The IRP stipulates the distribution of new capacity required from various energy sources up to 2030, with 17 800 MW allocated to RE (mainly solar photovoltaic (PV) and wind energy) [7]. Of this renewable capacity, 5000 MW must be operational by 2019 with an additional 2000 MW online by 2020. It was these requirements, along with the immediate need to extend generation capacity, that laid the foundation for the rapid and continuing development of RE generation in South Africa.

As reported by Eberhard *et al.* [2], the South African Department of Energy (DoE) first considered the use of renewable energy feed-in tariffs (REFITs) to encourage private investment and accelerate development of the renewable energy sector to the levels stipulated in the IRP. This approach, however, was later rejected in favour of competitive tenders, leading to the formation of the Renewable Energy Independent Power Producer Procurement Program (REIPPPP) in August 2011. The REIPPPP was established with the multi-faceted aim of expanding and diversifying the South African electricity generation capacity, decreasing its dependence on fossil fuels and enhancing the local renewable energy industry with regard to development, contracting, manufacturing and enhancement of skills [8]. In addition, it also intends to contribute toward local socio-economic development and environmentally sustainable growth.

The REIPPPP supports the procurement of electricity produced by the private sector using a variety of renewable technologies, namely onshore wind, concentrated solar power (CSP), solar PV, small hydro and biomass, which includes biogas, landfill gas and co-generation with agricultural waste or by-products. In order to aid the implementation of the IRP and the government's National Development Plan (NDP) 2030, which stipulates the addition of 10000 MW of electricity generation capacity to the 2013 baseline of 44 000 MW by 2020

Thus far, the implementation of the REIPPPP has been quite successful, featuring a well-organised bidding process, transparent and swift evaluation, and timely financial closure and construction of selected projects [2]. It has resulted in South Africa ranking among the top ten countries globally for renewable energy independent power producer (IPP) investment since 2012, as well as placing it among the top ten countries globally with regard to established utility-scale solar power capacity [9].

By December 2015 the REIPPPP had awarded a total of 5037 MW of renewable generation capacity to 77 projects over four rounds of competitive bidding, of which 1884 MW is allocated to a total of 29 solar PV installations [10, 11]. Table 1.1 presents a summary of the total generation capacity allocated throughout bid windows 1 to 4. The geographical distribution and relative generation capacity of all solar PV projects approved in these four bid windows are visually presented in Figure 1.1. Although there is a fair degree of overall geographical diversity in this collection of solar PV plants, it should be noted that more than 50 % of the aggregated generation capacity is confined to the indicated circular region in the center of the country.

With regard to the generation costs associated with the REIPPPP projects, the impact appears to be favourable in the context of the existing power generation network. Despite starting undesirably high in round 1, the bidding tariffs for the various RE sources have dropped significantly during each round as the competition was strengthened by increasing numbers of international private sector investors, and are now approaching equivalence to the tariffs for new coal-generated electricity [8]. This reflects international trends indicating that the costs of RE generation are becoming increasingly competitive with that of conventional fossil fuel generation options [12, 13].

Table 1.1: Summary of REIPPPP allocation of generation capacity [10, 11].

Renewable technology	**Allocation [MW]**				
	Window 1	***Window 2***	***Window 3***	***Window 4***	***Total***
Solar PV	632	417	435	415	1899
Wind	634	563	787	676.4	2660.4
CSP	150	50	200	0	400
Small hydro	0	14	0	4.7	18.7
Landfill gas	0	0	18	0	18
Biomass	0	0	16	25	41
Biogas	0	0	0	0	0
Total	1416	1044	1456	1121.1	5037.1

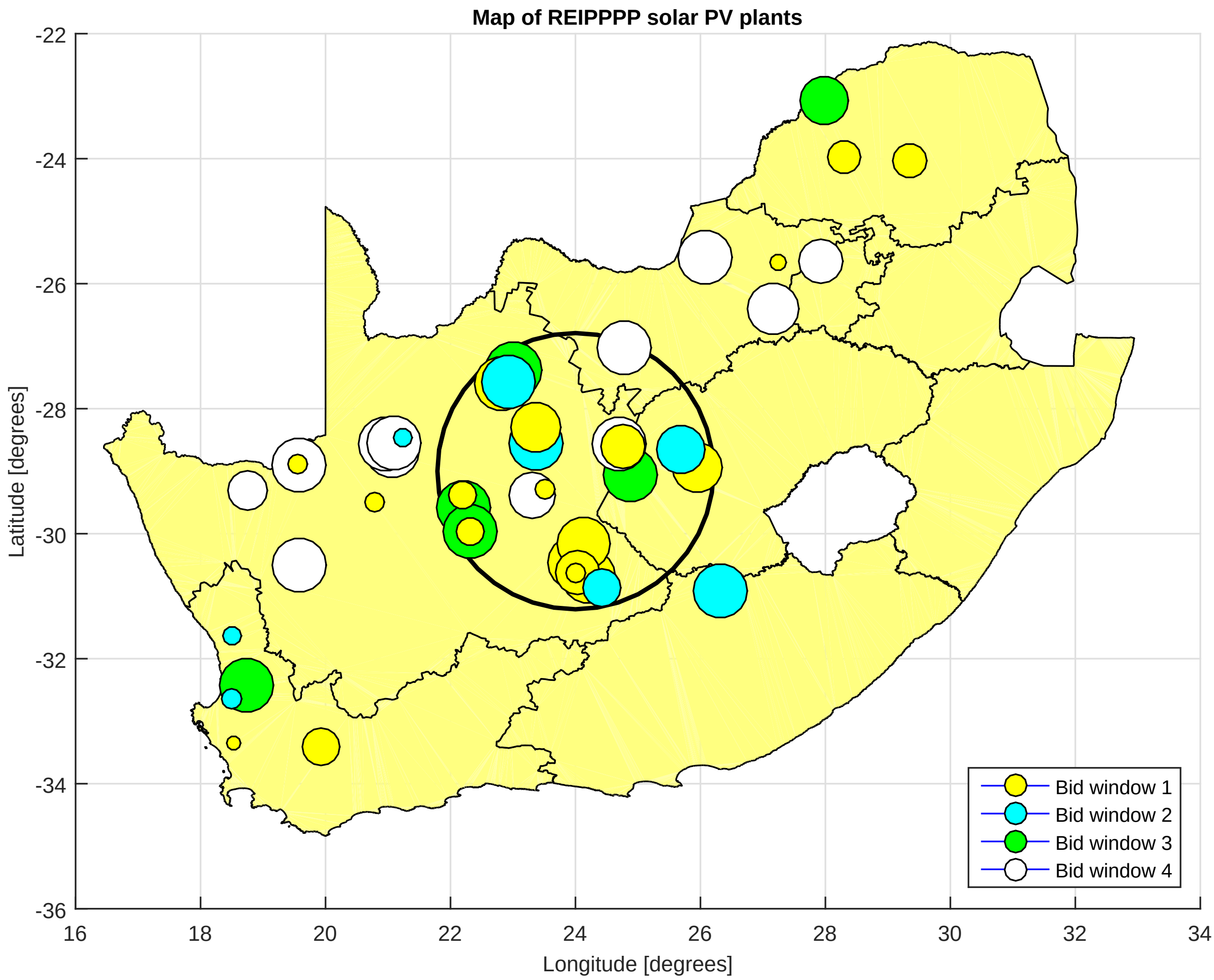

Figure 1.1: Map of the locations of all selected solar PV bids for the REIPPPP.

1.2 Challenges for large-scale grid-integration of RE

Despite offering economic viability and certain environmental advantages, the intermittent and variable nature of widely-used renewable technologies such as wind and solar PV systems can be problematic for both grid stability and supply reliability, especially for systems with high penetration of RE [14–16]. While the extent of this variability is generally mild or slow enough to be managed within a traditional grid context, much more extreme impacts may be experienced on occasions when wind farms or solar PV plants of several hundred mega-watts fluctuate unexpectedly between full and very low output in a matter of hours [17]. High RE penetration also affects cost-efficiency since renewable plants tend to be operated at a lower capacity factor as more capacity is installed, while significant renewable generation contributions result in increased generation reserve requirements [18]. In the case of wind generation, for example, the variability of generated power is absorbed within the variability of load at low penetration levels [17], while increased penetration can cause greater ramp-rates, inter-hour variability and scheduling errors, leading to greater reserve requirements [19].

In the local context, Bello *et al.* [20] described the technical implications for distribution network planning associated with increasing RE penetration in the local grid to the level specified by the IRP. The planned increase in RE distributed generation (DG) may lead to phenomena such as step voltage fluctuations due to inrush currents, unexpected voltage reduction due to generator disconnection, voltage rise due to reverse power flow from renewable plants, increased load on individual elements in the network, rapid voltage change (RVC) due to large changes in generation output and significant changes in network losses due to altered network configurations. Many of these considerations stem from the fact that the local sub-transmission and distribution networks were historically designed to operate radially, often with low levels of interconnectivity in rural areas, in order to reduce fault levels and simplify protection strategies. With the introduction of the REIPPPP, however, the number of generators connected to Eskom's transmission and distribution network, as well as the range of geographically dispersed connection locations, are set to increase significantly up to 2030.

Around the globe, the myriad of challenges experienced or anticipated at high RE penetration levels have necessitated the development of innovative grid management strategies, including advanced demand-side management, active network management via SMART Grid systems and optimisation of renewable energy resources with the view to improve grid support [21–23]. A substantial body of research indicates that a diverse portfolio of renewable generation capacity offers significant advantages over relying predominantly on one renewable resource [24–26]. Sovacool [27] found that deploying more renewable plants using a combination of resources, particularly wind and solar, increases grid stability and supply reliability. As for the impact of RE on generation reserve requirements, Halamay *et al.* [19] found that the diversification of renewable generation capacity with a combination of wind, solar and/or wave energy also consistently improves the reserve requirements compared to the wind-alone case.

European countries have benefited from their extensively interconnected national grids in handling the complexity associated with grid-integrated RE generation, since it facilitates convenient exportation of surplus energy to areas of the network where the current demand is high [18]. South Africa with its relatively isolated national grid, however, needs to design alternative solutions to maintain network integrity and achieve economic dispatch with high penetration renewable power generation. One mitigation option for handling intermittent renewable energy supply is pumped storage schemes, which has already been implemented locally in the established Steenbras, Palmiet and Drakensberg pumped storage schemes, rated at 180 MW, 400 MW and 1000 MW respectively [20]. In addition, the new Ingula scheme, rated at 1333 MW, is due to become operational in 2017, with the 1520 MW Tubatse scheme scheduled for 2022 [28]. Incorporating mitigation measures, however, should be secondary to developing a diversified generation system that is optimised and self-sufficient due to thorough long-term analysis and planning.

1.3 Optimising renewable energy generation

Appropriate site selection for non-dispatchable renewable energy sources is an integral component of long-term energy system development and grid integration planning. The research reported in literature for determining optimal solar PV plant locations generally focus on network constraints, the availability of grid connection infrastructure and individual plant profitability based on maximum annual energy generation. Gomez *et al.* [29] investigated the optimisation of PV plant location by means of a particle swarm methodology, using cost and profitability parameters as inputs with a flat annual feed-in tariff. Huy *et al.* [30] used Differential Evolution (DE) to optimise the location, sizing and power factor of solar PV plants for sequential or concurrent integration into the grid such that network loss is minimised and power system constraints are not violated. Urquhart *et al.* [31] conducted a multi-objective optimization study to allocate a fixed PV capacity among a selection of candidate sites such that energy is maximised while variability in the form of ramp rates is minimised according to a specified design point. Carrin *et al.* [32] combined multi-criteria analysis, the analytic hierarchy process and geographical information systems to define a decision-making model that incorporates a wide selection of environmental, orographic, location-based and climatological indicators. In the local context, a comparative analysis of solar PV plant location scenarios has been conducted for South Africa, with the focus on grid infrastructure, the cost of transmission losses and requirements for distribution and transmission system upgrades [33].

From the perspective of grid support, the incorporation of load considerations into the optimisation process has the potential to improve operational performance with reference to aspects such as availability and system stability, especially for systems with high penetration of renewable sources. This approach, in principle, allocates solar PV generation capacity in such a way that the aggregated PV power output profile provides improved grid support in the context of the diurnal and seasonal characteristics associated with the system load profile, thereby introducing the important operational aspect of economic dispatch. By improving the availability of renewable energy generation during peak demand periods, the use of expensive peak generation plants can be mitigated [27]. Many developing economies experience capacity constraints during peak periods, giving rise to high peak generation costs due to the use of expensive generation sources .

Locating renewable energy sources for maximum annual energy yield can result in many large installations located geographically in close proximity, which for solar PV plants typically translates to the region with the highest solar radiation levels. The daily and seasonal power generation profiles associated with this scenario yields high mean values, but can also yield high variability for the aggregated power generation cycle. Furthermore, local weather conditions can give rise to severe loss of the renewable generation component, both for solar PV and other sources such as wind. Spatial diversification of renewable generation, however, has proved useful in decreasing variability and facilitating power flow [27]. Finding the optimal implementation of this diversification then becomes key to effective expansion and integration of the renewable generation fleet.

Wind power generation, with its dependence on stochastically variable wind speed and wind, is notoriously difficult to forecast, even with extensive historical data, and will always have a significant margin of error [34]. Solar PV power generation, which is largely dependant on solar radiation, is less uncertain although it can be significantly affected by location-specific weather fluctuations in the short-term [19]. From a long-term perspective, however, the availability of solar energy produces clear seasonal trends determined by the solar cycle as well as climatological factors [35]. A significant amount of research has been conducted to establish accurate methods of short-term power output prediction for solar PV plants using various environmental factors, as discussed in a review by Mellit *et al.* [36]. These methods are commonly utilised by grid-integrated PV solar plants as tools for predicting and optimising their day-to-day energy dispatch scenarios.

From a long-term grid-support perspective, however, the day-to-day operation of individual plants is of less concern than the aggregated yield forecast of all PV solar plants averaged over specific months and seasons, along with the statistical probability of achieving these averaged forecasts for a given day and time. Several significant long-term factors affecting the performance of a solar PV plant in this regard can be clearly identified, namely the longitude (which determines the diurnal cycle), the latitude (which determines the seasonal cycle) and the climate (which determines the amount of solar energy available in the long-term as well as the impact of adverse weather conditions). South Africa is a good candidate for investigating the impacts of these factors as it has significant variation in longitude and latitude with strong seasonal weather variations along both the east-west and north-south axes. Similar arguments apply for large interconnected power systems spanning large geographical areas.

1.4 Project description

1.4.1 Overview

Based on the growing energy demand and increasing significance of the RE sector in South Africa, there is a clear need for effective methodologies and tools for planning the expansion of power generation capacity, particularly when utilising RE sources. It follows that grid-integration of the RE fleet would be most beneficial if the aggregated long-term energy generation profile of the distributed RE plants were optimised to satisfy known load characteristics. In view of this consideration, this project explores one such optimisation methodology via a comprehensive optimisation study that investigates the allocation of solar PV generation capacity in the context of various grid-support metrics.

1.4.2 Project scope

The optimisation study involves the design of an optimisation strategy based on meaningful parameters pertaining to both solar PV power generation and long-term demand considerations in the South Africa context, as well as the subsequent evaluation and comparative analysis of the designated optimisation problem cases. The optimisation strategy comprises the design of problem cases based on relevant aspects of the diurnal and seasonal cycles for both the solar resource and the national demand profile, along with the identification of suitable grid-support optimisation objectives and optimisation algorithms for evaluating the resulting problem cases.

The project specification requires that the optimisation study be implemented within an integrated software platform that is to form part of an ongoing software project. The software design element of the project encompasses the implementation of a solar PV optimisation module as a plug-in application within an established database-driven user interface, along with the integration of the necessary optimisation simulation software. While certain aspects of the software implementation are prescribed due to the objective of integration with an ongoing project, the structural design of the relational database as well as the functionality of the solar PV optimisation module and its simulation software require due consideration.

As this is an investigative study aimed at assessing the potential contribution of selected methodologies and techniques in facilitating energy generation planning, the scope of the project does not extend to attaining concrete, specific solutions for immediate practical application. The intent is rather to create a useful foundation, in terms of both the implementation of the integrated software platform, as well as analysis of the optimisation study results to identify underlying trends and considerations pertinent to future work. To this end, the software design aspect of the project is approached with particular regard for potential expansion and addition of features. Generic infrastructures for both the database and the user interface are desired in order to accommodate future studies using energy sources other than solar PV and/or different analytical methods.

1.4.3 Project objectives

The project objectives can be categorised as follows:

- Design objectives pertaining to the implementation of the integrated software platform.
- Research objectives pertaining to the various optimisation algorithms and problem cases considered in the optimisation study.

The following design objectives were identified with regard to the implementation of the integrated software platform:

- A highly organised database structure capable of handling large data sets and a wide variety of data profiles and simulation parameters efficiently.
- A modular approach to the implementation of the solar PV optimisation module as well as optimisation simulation software in order to facilitate future expansion of the software functionality.
- A generic approach to the structural design of the relational database as well as the optimisation module with the view to accommodate different types of data and energy sources in future studies.

The following research objectives were then identified with regard to the optimisation study:

- Investigation of the long-term implications of the diurnal and seasonal cycles pertaining to both solar and demand profiles for optimising the allocation of solar PV generation capacity.
- Investigation of the effects of optimising the allocation of solar PV generation capacity for various grid-support objectives.
- Performance analysis of the selected optimisation techniques with regard to evaluating the type of search problems.

1.4.4 Key questions

The following key questions pertaining to the software design component of the project were identified:

- Which platform would be most suitable for performing the required optimisation simulations?
- How will the various system components be effectively integrated into a single platform?

- How can the optimisation module and database structure accommodate a wide range of potential future applications?
- How can scalability of data and expandability of functionality be inherently included in the system design?

Finally, the following key questions pertaining to the specified research objectives were identified:

- What type of data would be suitable to use for solar PV performance evaluation and how can that data be sourced?
- Which grid-support characteristics are most significant for RE integration and in what format can they be used as optimisation parameters or objective functions?
- What additional optimisation parameters are important in the context of optimising the allocation of solar PV generation capacity?
- Which optimisation algorithms would be suitable for analysing the selected optimisation problem cases and how can they be implemented?

1.4.5 Project tasks

The aforementioned project objectives and key questions translate into the following research and design activities:

- Conducting a thorough literature study to achieve a full understanding of the local power generation and renewable energy sector, as well as all relevant solar theory, software tools, analytical techniques and relevant past research.
- Evaluating and optimising the existing database topology of the ongoing software project for organised storage of a wide range of plant and optimisation parameters, measured and simulated power profiles, and analytical result sets.
- Developing a suitable user interface for the required solar PV optimisation module and implementing efficient integration with the established application.
- Identifying a suitable set of objective functions and other optimisation parameters derived from relevant grid-support metrics as well as seasonal and diurnal considerations.
- Identifying and obtaining suitable input datasets for evaluating solar PV performance in the context of this study, and populating the database with this suite of data.
- Identifying suitable optimisation algorithms and implementing these methods in simulation software integrated with the solar PV optimisation module.

- Defining a set of optimisation problem cases based on the selected input data, objective functions, optimisation parameters and optimisation techniques.
- Evaluating the specified optimisation problem cases via the integrated simulation platform and analysing the resulting solutions in the context of the research objectives.

1.5 Thesis overview

The contents of this document is structured as follows:

- *Chapter 1:* Detailed discussion of the project motivation and project description, including the project scope, research and design objectives, key questions and associated project tasks.
- *Chapter 2:* Deconstruction of the optimisation strategy and design of optimisation problem cases, with reference to seasonal and diurnal considerations, optimisation objectives identified with regard to grid-support metrics, and optimisation algorithms selected for the evaluation of problem cases.
- *Chapter 3:* Discussion of the implementation of the optimisation study in the context of the integrated software platform, including design and implementation of the required software infrastructure, acquisition of suitable solar power profiles, and realisation of appropriate objective functions for evaluating the specified optimisation objectives.
- *Chapter 4:* Presentation and analytical discussion of the results obtained for the various problem cases evaluated in the context of the designated optimisation study.
- *Chapter 5:* Discussion of conclusions and recommendations for future work derived in the context of the optimisation study implementation and results.

Chapter 2

Optimisation strategy and problem case selection

2.1 Overview

As described in Section 1.4, the main research objective of this project is the investigation of strategies for optimising solar photovoltaic (PV) plant locations to provide grid support in a long-term capacity. Considerations for grid support include both reliability in terms of the daily variation presented by solar PV power generation, as well as the extent of its contribution during peak demand periods. When considering solar PV performance from a long-term perspective, the role of seasonal and diurnal cycles emerges as highly significant to the effort of improving grid support. These cyclical parameters are pertinent both in terms of the regional location of solar PV installations, as well as their performance throughout specific periods in those cycles.

With regard to the solar resource that solar PV power generation depends on, the presence and significance of both seasonal and diurnal cycles are quite obvious. Furthermore, there are strong correlations between this diurnal cycle and the geographical east-west axis as well the seasonal cycle and the north-south axis. Eskom's energy tariff structures, in particular the Megaflex tariff structure that applies to the majority of their industrial and municipal reseller clients, also present clear seasonal and diurnal cycles [37]. In this context, the diurnal cycle is reflected in the daily Time-of-Use (TOU) schedule for which the weekday specification is presented in Table 2.1. Meanwhile, the seasonal cycle consists of a high demand and low demand season as shown in Table 2.2. Since this tariff structure is informed by the national demand profile and the corresponding cost of generation [38], it is a highly relevant consideration for energy generation optimisation.

Table 2.1: Eskom's Megaflex TOU schedule for weekdays [37].

Time period	TOU period
06:00 - 07:00	Standard
07:00 - 10:00	Peak
10:00 - 18:00	Standard
18:00 - 20:00	Peak
20:00 - 22:00	Standard
22:00 - 06:00	Off peak

Table 2.2: Eskom's Megaflex tariff structure for 2014/2015 [37].

TOU period	Energy charge [ZAR/kWh]	
	High demand season	***Low demand season***
Peak	2.2384	0.7302
Standard	0.6780	0.5025
Off-peak	0.3682	0.3189

In view of this two-fold importance, the selection of the potential solar PV plant locations considered in this investigation was based on the aforementioned seasonal and diurnal considerations. Furthermore, these cycles were also incorporated as independent optimisation parameters. A selection of optimisation objectives aligned with various grid support aspects were identified for evaluation, along with several optimisation algorithms for performing the required optimisations. Consequently, this optimisation study consists of a range of problem cases composed of the following parameters: a selection of potential plant locations, a diurnal and seasonal specification, an optimisation objective, and optimisation algorithms.

In accordance with the main analytical objective, each problem case was evaluated by optimising the allocation of a normalised per-unit generation capacity (i.e. a generation capacity of arbitrary size) within the given selection of potential locations according to the relevant optimisation objectives and cyclical parameters. Since the selected optimisation objectives all concern some aspect of solar PV performance in terms of either power or energy output, the solar power profiles for potential locations were used as the sole performance indicator for all problem cases.

2.2 Seasonal and diurnal cycles

2.2.1 Seasonal cycle

Due to the Earth's tilted axis and the nature of its yearly orbit around the sun, there is a strong seasonal cycle associated with the availability of solar energy at any given location [39]. Figure 2.1 presents the estimated north-facing sun path diagram ranging from 21 December (i.e. the summer solstice) to 21 June (i.e. the winter solstice) for the city of Polokwane, which is located towards the northern extremity of South Africa. The diagram shows a significant variation in daylight hours as well as the elevation of the sun relative to the horizon throughout

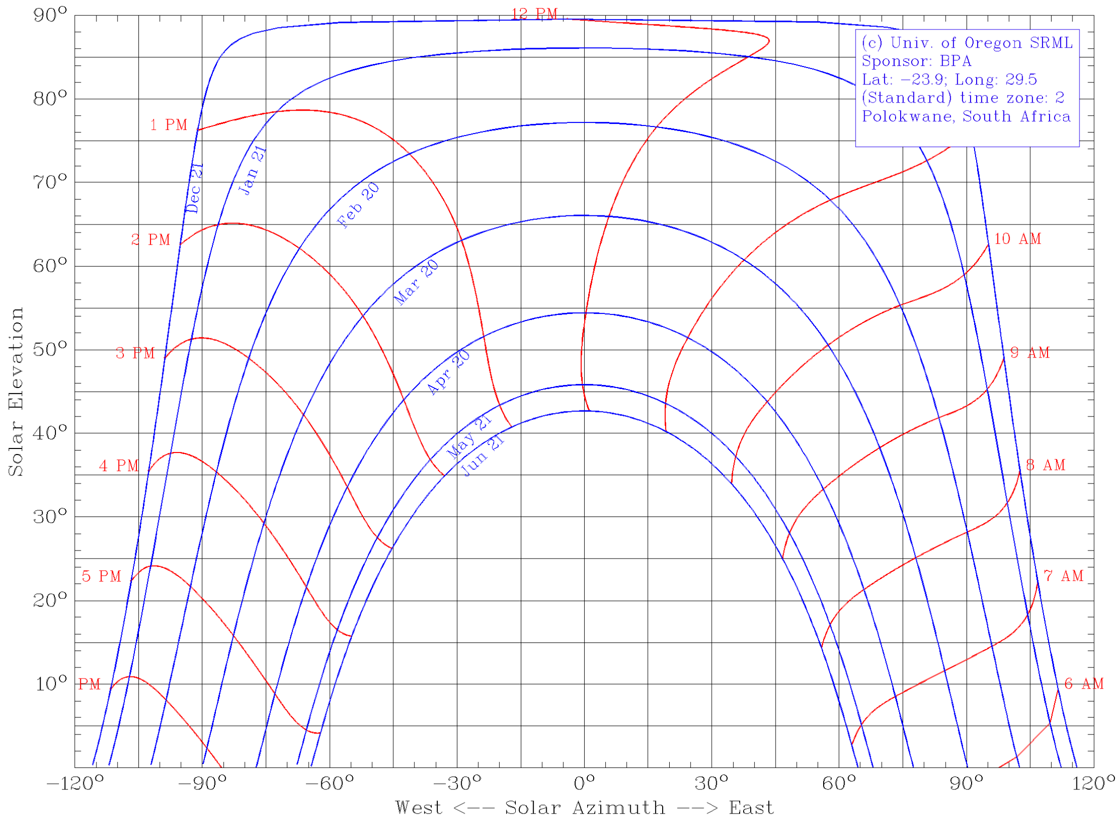

Figure 2.1: Estimated variation of solar path observed in Polokwane between 21 December and 21 June [40].

the year. The amount of daylight as well as the solar elevation range from their highest level at the summer solstice to their lowest level at the winter solstice.

Figure 2.2 presents the corresponding sun path diagram for the city of Port Elizabeth, which is situated near the southern extremity of the country. While also exhibiting a notably seasonal pattern, the form of this diagram clearly varies from the solar path observed at Polokwane, which affirms the crucial role of regional latitude with regard to the seasonal availability of solar energy. Conversely, Figure 2.3 presents the sun path diagram for a south-facing observer in Frankfurt, Germany, which illustrates the reversal of the seasonal solstices between the southern and northern hemisphere. The extreme latitude of this location relative to the South African locations is also evident in the form and seasonal variation of its sun path diagram. The seasonal cycle for the solar resource as well the corresponding variation along the north-south axis is thus clearly established.

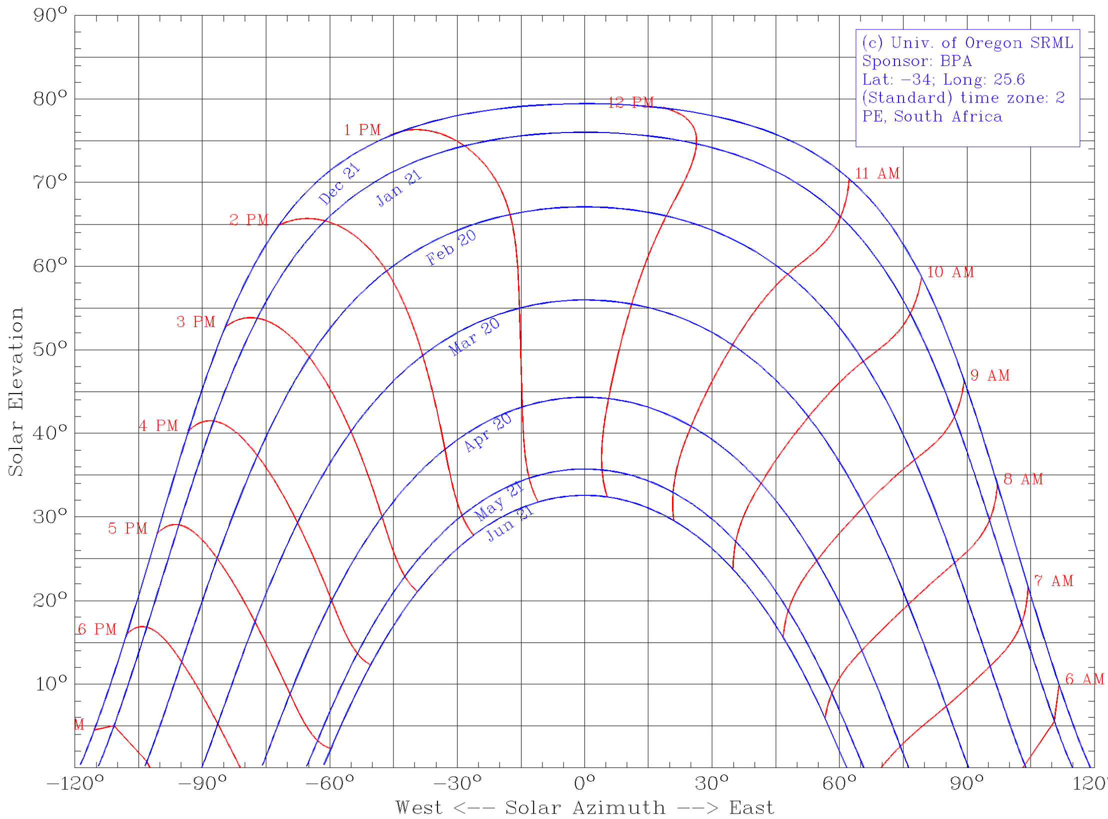

Figure 2.2: Estimated variation of solar path observed in Port Elizabeth between 21 December and 21 June [40].

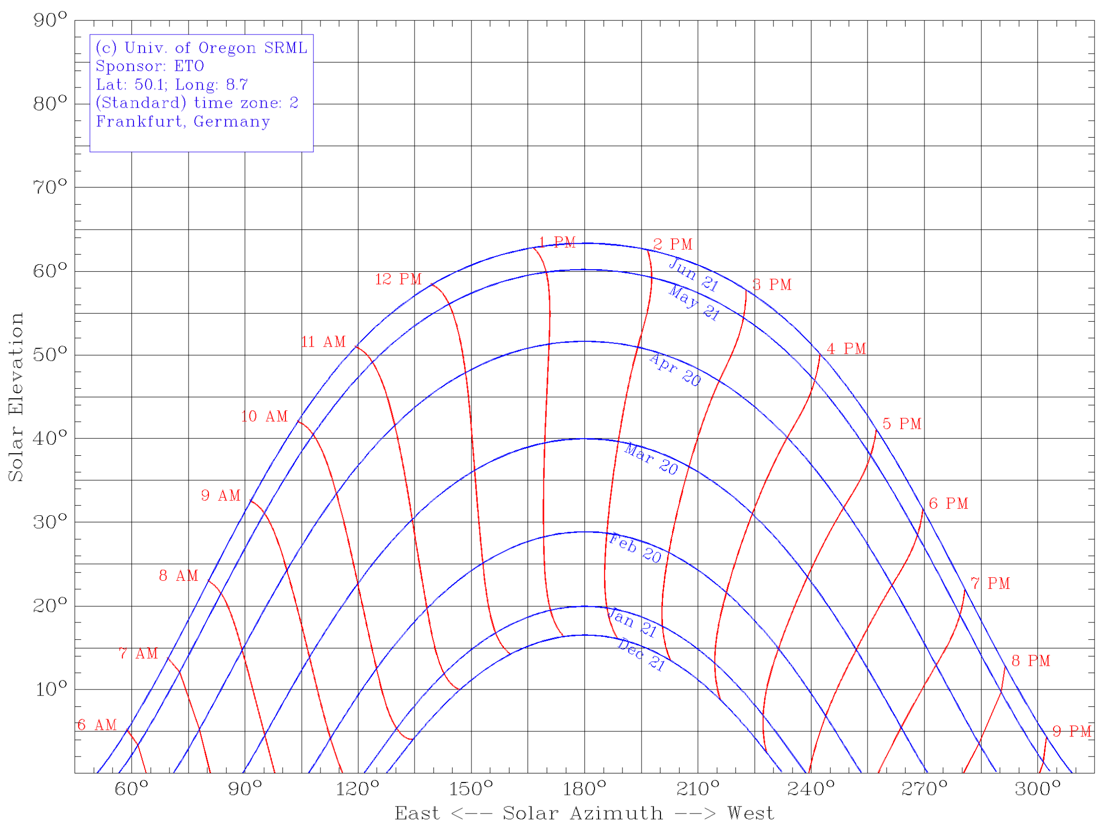

Figure 2.3: Estimated variation of solar path observed in Frankfurt between 21 December and 21 June [40].

Apart from the influence of the seasonal solar cycle, the long-term regional availability of solar energy is also characterised by climatological patterns. In this regard, the South African climate exhibit strong seasonal trends along the north-south axis. The northern part of the country represents a summer rainfall region, where the summer season is associated with hot and humid conditions and thunderstorm activity in the afternoons, while the winter season is associated with dry and cold conditions with clear skies. Meanwhile, the southern part of the country represents a winter rainfall region, where the summer season is associated with dry and hot conditions with clear skies, while the winter season is associated with wet and cold conditions with cloudy skies. In addition to these attributes, significant climatological features can also be distinguished along the east-west axis. This is due to the effects of the cold Benguela current travelling north in the Atlantic ocean on the western side of the country and the warm Agulhas current travelling south in the Indian ocean on the eastern side of the country. The result is that the western part of the country experiences a semi-desert climate, characterised by dry and hot conditions and clear skies for most of the year. Conversely, the eastern part of the country experiences a subtropical climate characterised by humid and hot conditions with hazy skies and thunderclouds.

In view of this cyclical behaviour, the optimisation of solar performance for individual seasons as well as the full year scenario is investigated for the various optimisation objectives. The months allocated to each season are as follows:

- *Summer*: December, January and February
- *Autumn*: March, April and May
- *Winter*: June, July and August
- *Spring*: September, October and November

With regard to the seasonal cycle of the grid, Eskom defines an annual low demand and high demand season for its Megaflex tariff structure that includes a distinct pricing adjustment, as shown in Table 2.2. The high demand season consists of the months of June, July and August, i.e. the traditional winter months, while the low demand season extends from September through to May [37]. Consequently, the results of problem cases optimised for the winter months is of particular interest throughout this study, especially compared to the corresponding results for problem cases optimised on an annual basis.

While its merit in terms of grid support is evident, optimisation with regard to the seasonal demand cycle may also be pertinent to the profitability of grid-connected solar PV plants. Should South African renewable energy (RE) policy-makers replace the flat feed-in tariffs that currently apply to large-scale PV installations with variable feed-in tariffs that reflect the demand profile and current cost of generation, plants could generate more revenue by maximising the energy delivered during the high demand season rather than the energy delivered annually.

2.2.2 Diurnal cycle

The nature of the solar diurnal cycle is self-evident and its correlation with location in terms of the east-west axis is intuitive. Figures 2.4 and 2.5 present the sun path diagrams associated with the South African locations of Alexander Bay and Durban, respectively, which are situated at the longitudinal extremities of the country. These diagrams are highly similar with regard to form and total sunlight hours, yet there is a distinct time shift that shows earlier sunsets and sunrises for Durban, the eastern location. This behaviour confirms the significance of regional longitude for solar PV performance.

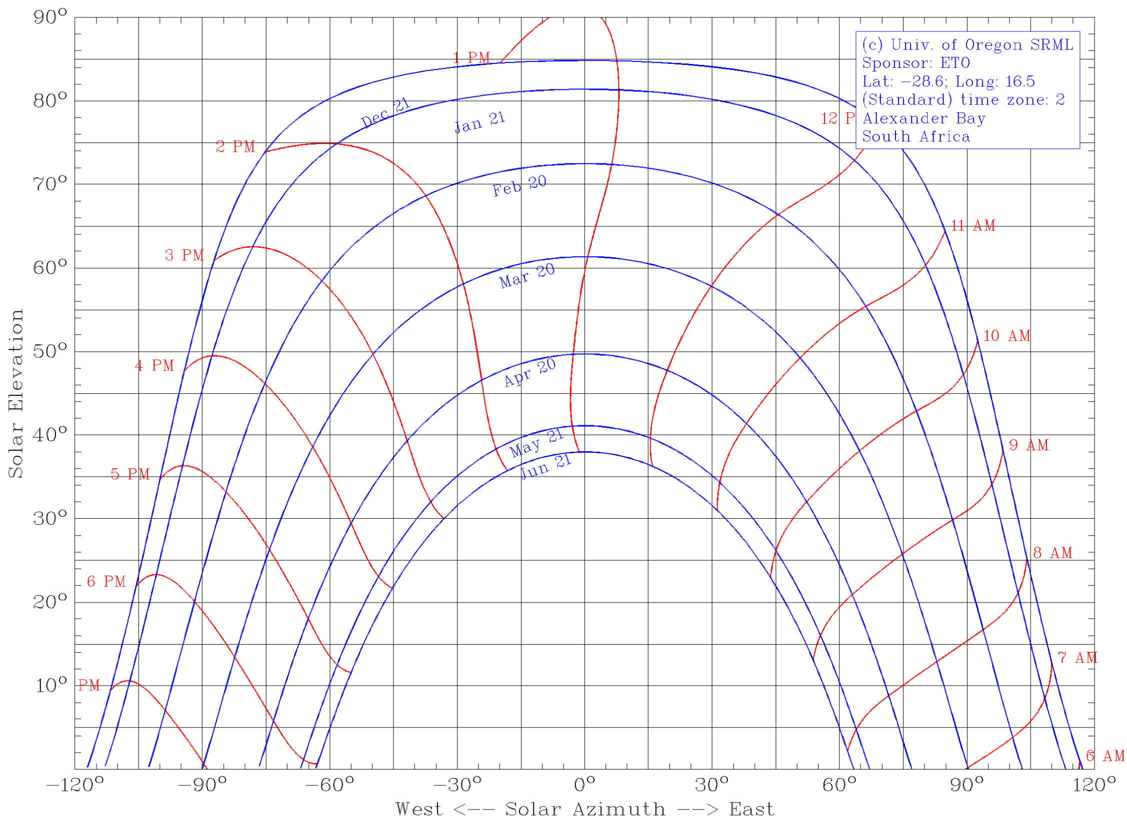

Figure 2.4: Estimated variation of solar path observed in Alexander Bay between 21 December and 21 June [40].

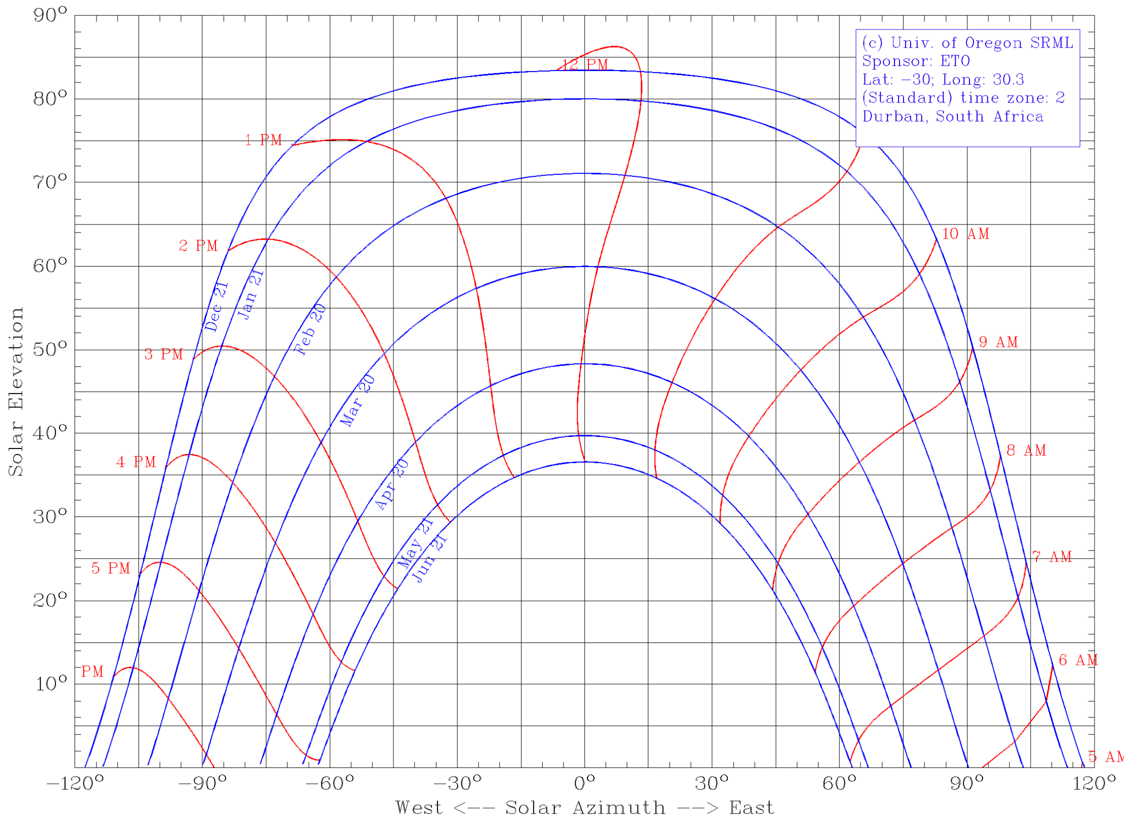

Figure 2.5: Estimated variation of solar path observed in Durban between 21 December and 21 June [40].

In the context of the grid, the daily TOU schedule for Eskom's Megaflex tariff structure is a reflection of the national diurnal demand profile as well as the corresponding generation costs [38]. This schedule identifies peak, standard and off-peak TOU periods and varies for weekdays, Saturdays and Sundays [37]. The weekday schedule, shown in Table 2.1, includes two peak tariff periods while the weekend schedules consist only of standard and/or off-peak charges. Since it is not expedient to optimise solar PV generation with different schedules for different days, only the weekday schedule was considered for the purposes of this study. Based on this schedule, the following diurnal specifications were identified as meaningful optimisation parameters for comparative analysis:

- Full day scenario (i.e. all sunlight hours)
- Morning peak period (07:00 - 10:00)
- Evening peak period (18:00 - 20:00)
- Combined morning and evening peak periods

Due to the higher national load and associated generation costs per kilowatt-hour (kWh) implied by these intervals, the benefits of maximising power generation over the peak periods might potentially be greater than simply optimising for all sunlight hours. In terms of the demand profile, such an approach could serve to alleviate the strain on the grid during high demand periods, thereby reducing the operation of costly diesel generators which currently supplement generation during peak demand periods. As in the case of the seasonal demand cycle, optimisation from the diurnal perspective could also be highly advantageous in terms of plant revenue if a variable feed-in tariff should be introduced for grid-connected solar PV installations.

It should be noted, however, that even at the western extremity of South Africa the sunset times are relatively early compared to the evening peak period, ranging from 18:05 at the winter solstice to 19:51 at the summer solstice. Conversely, the sunrise times at the eastern extremity, which range from 06:51 at the winter solstice to 04:52 at the summer solstice, produces a much more significant amount of sunlight during the morning peak period. From these specifics it can be deduced that solar PV power generation could make a significant contribution during the morning peak period, but is not a useful power source for the evening peak period in South Africa. Optimisation with regard to the evening peak still merits evaluation, however, as it may offer more significant benefits for networks spread over larger geographical regions or subject to greater seasonal variation.

2.3 Site groups

All potential solar PV plant locations considered in this optimisation study were selected specifically to investigate the seasonal and diurnal considerations previously discussed. The potential locations were also restricted to the South African locations for which solar data is available from chosen meteorological data source. Based on this limitation, three groups of locations (forthwith referred to as site groups) were selected for evaluation, as detailed in Table 2.3. The locations in site groups 1 and 2 were chosen to investigate, respectively, the impacts of the seasonal and diurnal solar cycles on potential solar PV power generation, while site group 3 combines all the locations in site groups 1 and 2 for a more general approach. Comparative analysis of the optimisation results for these three site groups therefore offers a thorough representation of the seasonal and diurnal trends, while also illustrating the effects of optimising for a larger number of potential locations with wider geographical dispersion.

Table 2.3: Site group locations.

Location	Abbreviation	Latitude	Longitude	Site group		
				1	2	3
Alexander Bay	AB	-28.6	16.5	✓		✓
Bloemfontein	BFN	-29.1	26.3	✓	✓	✓
Durban	DBN	-30	30.9	✓		✓
Kimberley	KIM	-28.8	24.8	✓		✓
Middelburg	MID	-31.5	25		✓	✓
Polokwane	PKW	-23.9	29.5		✓	✓
Port Elizabeth	PE	-34	25.6		✓	✓
Pretoria	PRE	-25.7	28.2		✓	✓
Upington	UTN	-28.4	21.3	✓		✓

In accordance with the longitudinal dependence of the solar diurnal cycle, the locations in group 1 are situated at similar latitudes but distributed across the breadth of South Africa along the east-west axis, as depicted in Figure 2.6. Meanwhile, Figure 2.7 shows that the locations in group 2 are of similar longitude but distributed along the north-south axis of the country to account for the seasonal solar cycle. Group 3 simply contains all locations in group 1 and 2, thus incorporating geographical variation along both the east-west and north-south axes.

Figure 2.6: Map of the locations in site group 1.

Figure 2.7: Map of the locations in site group 2.

2.4 Solar power

2.4.1 The solar resource

To a great extent, the maximum possible power output of a solar PV panel relative to its rated power is determined by the solar irradiance delivered to the panel at any given moment [41]. Solar irradiance, also known as solar radiation, denotes the instantaneous intensity of electromagnetic radiation from the sun that is normal/perpendicular to a collector surface (e.g. a PV panel) and is typically measured in W/m^2 [39]. The cumulative sum of solar irradiance collected on such a surface over a specified time interval is referred to as solar irradiation or insolation and is commonly measured in Wh/m^2.

The total solar radiation that strikes a collector surface consists of three elements, namely the component of the Direct Normal Irradiance (DNI) that is perpendicular to the surface, the diffuse sky radiation and the reflected ground radiation [42]. DNI refers to the irradiance on a surface normal to the direct beam of radiation delivered from the sun to a given location on the Earth's surface, as depicted in Figure 2.8. Diffuse sky radiation, as shown in Figure 2.9, includes incoming solar radiation that is scattered towards the ground by atmospheric particles and moisture, reflected downwards by clouds or reflected by the Earth's surface and then scattered back down [39]. Consequently, the diffuse radiation is usually a much more significant component of the total on hazy and cloudy days. Since precise estimation of diffuse sky radiation is notoriously difficult, it is often assumed to be isotropic, i.e. radiating equally from all directions, which means that it is proportional to the fraction of the sky visible to a collector surface [42].

Where diffuse radiation comes from the sky, reflected radiation results from solar radiation reflected by surfaces in front of the collector surface, as depicted in Figure 2.10. It is generally a small to negligible component of the total, with some models omitting it entirely, although nearby surfaces of a highly reflective nature, e.g. snow or large bodies of water, may increase it to a significant level [39]. Without such surfaces, the terrain surrounding a collector surface is assumed to reflect radiation with equal intensity in all directions, which renders reflected radiation proportional to the terrain that is visible to a collector surface. The reflective property of a given terrain can be described by its albedo, which is the portion of incoming radiation that is reflected by its overall surface [43]. Like diffuse sky radiation, reflected radiation is difficult to estimate accurately and modelling it necessitates certain assumptions; consequently, most models are likely to underestimate rather than overestimate both these components [42].

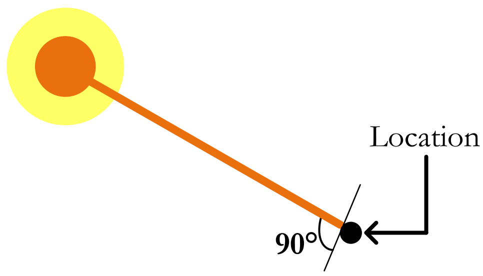

Figure 2.8: DNI at a given location (adapted from [39]).

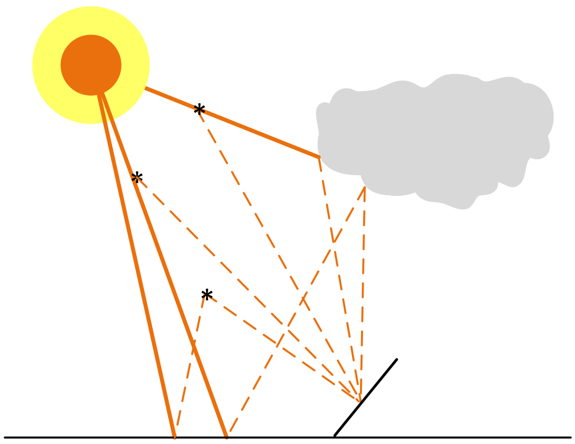

Figure 2.9: Diffuse radiation striking a collector surface (adapted from [39]).

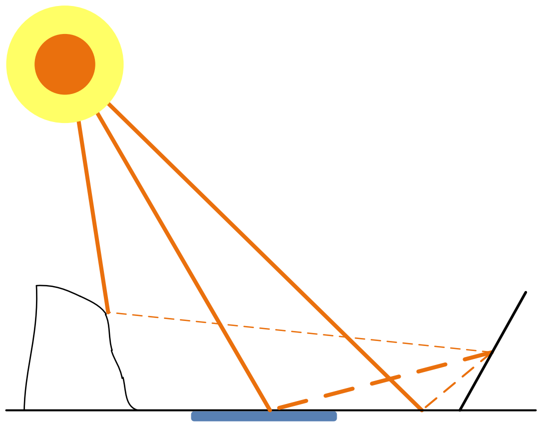

Figure 2.10: Reflected radiation striking a collector surface (adapted from [39]).

The component of DNI that is normal to a collector surface depends on the relative positions of the sun and the surface in question. As shown in Figure 2.11, the solar position at a given time and location can be described by an altitude angle β (also referred to as the solar elevation), which is measured relative to the Earth's surface, and an azimuth angle ϕ_S, which is measured relative to north in the southern hemisphere and relative to south in the northern hemisphere [39]. As indicated in Figure 2.12, the position of the collector surface is also described by an azimuth angle ϕ_C along with an elevation angle Σ, which is measured relative to the horizontal ground surface. From these relative positions the incidence angle of the DNI on the collector surface can be derived. The incidence angle θ is measured relative to the normal of the surface as illustrated in Figure 2.13, and describes the component of DNI that strikes the collector surface, denoted by I_{BC}, via the following formula [39, 42]:

$$I_{BC} = DNI \cdot \cos\theta \tag{2.1}$$

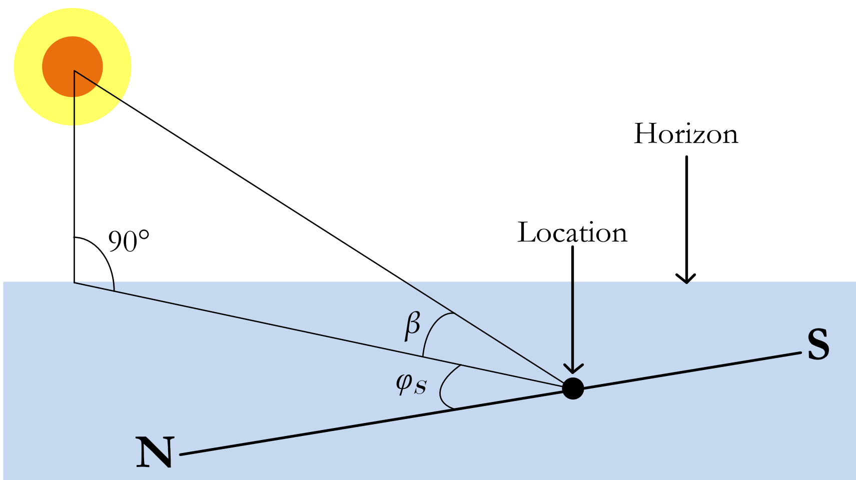

Figure 2.11: Description of solar position at a given location (adapted from [39]).

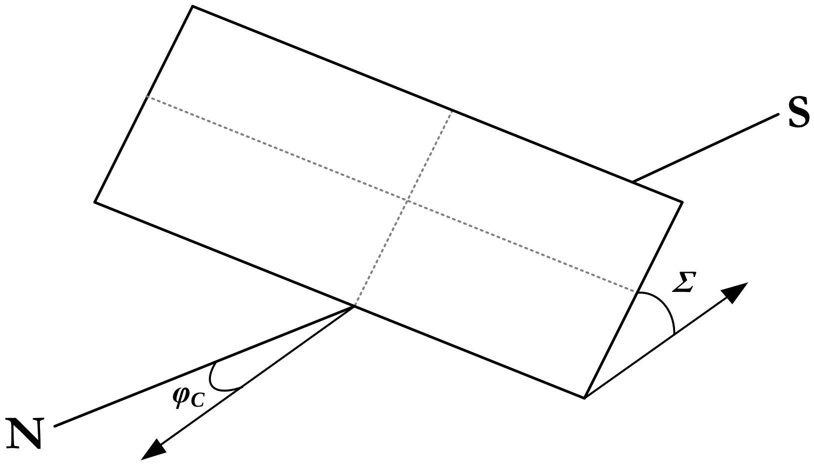

Figure 2.12: Description of collector surface position (adapted from [39]).

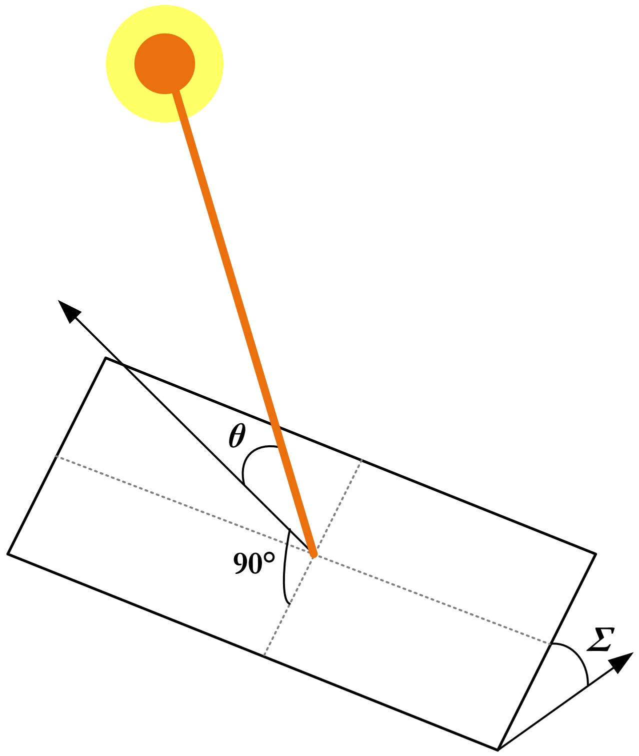

Figure 2.13: Incidence of DNI on a collector surface (adapted from [39]).

When the collector surface is taken as horizontal relative to the Earth's surface, the total perpendicular solar radiation is known as Global Horizontal Irradiance (GHI), while the total radiation normal to any tilted surface is known as Global Tilted Irradiance (GTI) [42].[1] Annual

[1]It should be noted that the acronyms DNI, GHI and GTI may be used to refer to either power (i.e. irradiance) or energy (i.e. irradiation). For the sake of clarity, however, they will only be used in the context of power profiles throughout the remainder of this document. References to irradiation will be explicitly indicated where necessary.

global horizontal irradiance, i.e. the cumulative solar energy delivered to a horizontal surface over the course of a year, is often used to evaluate the potential for solar PV power generation at a given location. Figures 2.14 to 2.17 present solar maps of the annual global horizontal irradiance as well as the direct normal irradiance throughout South Africa and the world as a whole [44]. A comparison of these maps indicate that the annual solar energy available for the majority of land area in South Africa is quite high by international standards, which makes it a very favourable candidate for large-scale solar PV power generation.

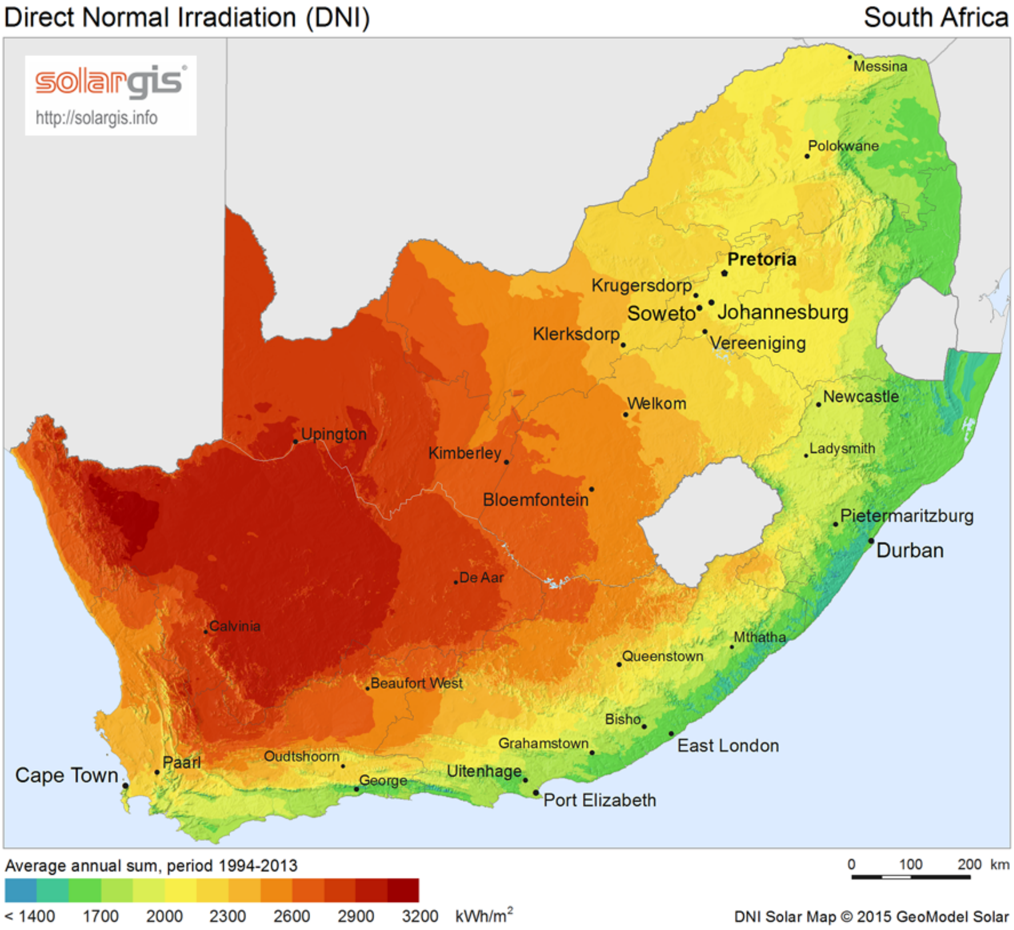

Figure 2.14: Annual direct normal irradiation throughout South Africa [44].

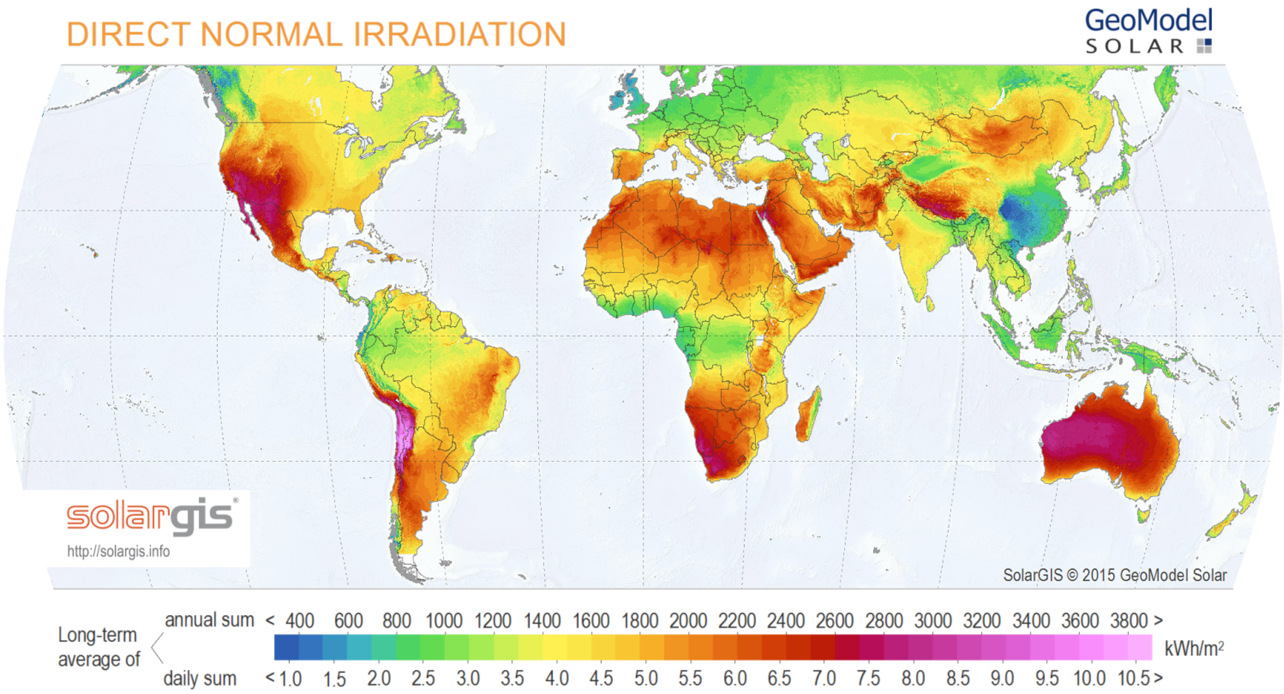

Figure 2.15: Annual direct normal irradiation throughout the world [44].

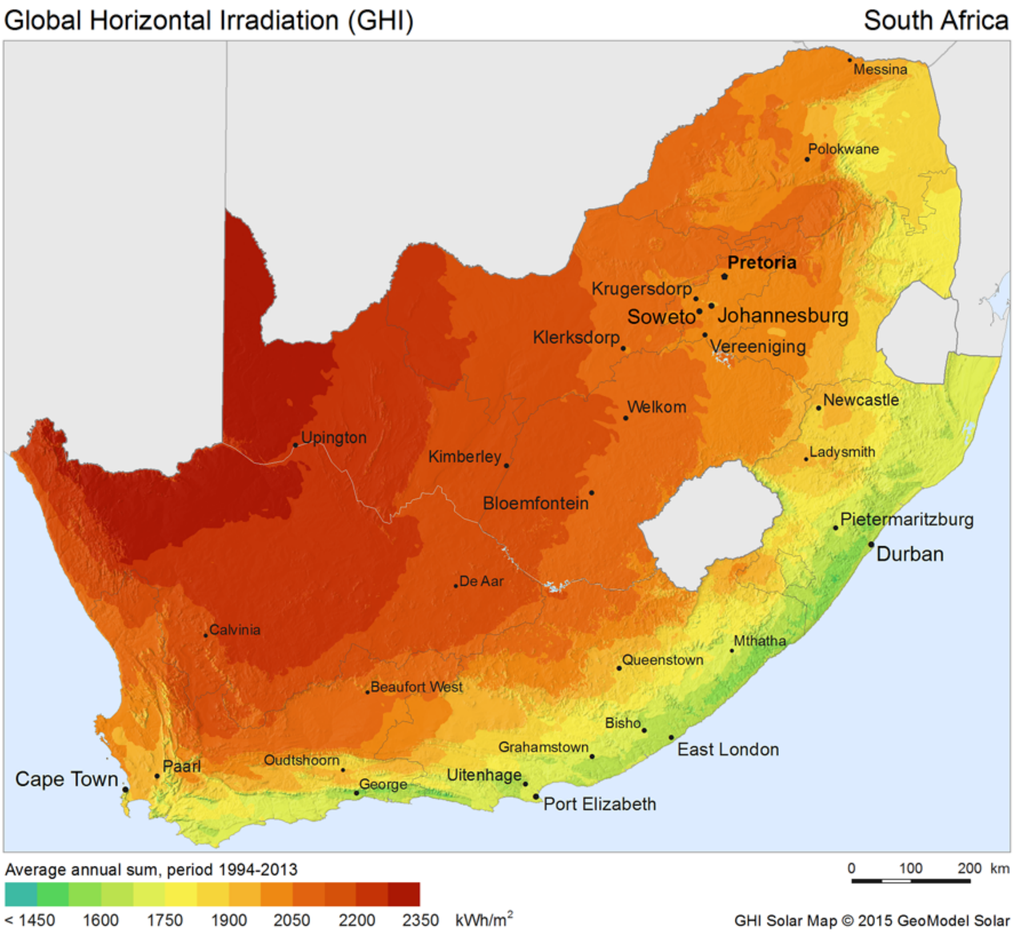

Figure 2.16: Annual global horizontal irradiation throughout South Africa [44].

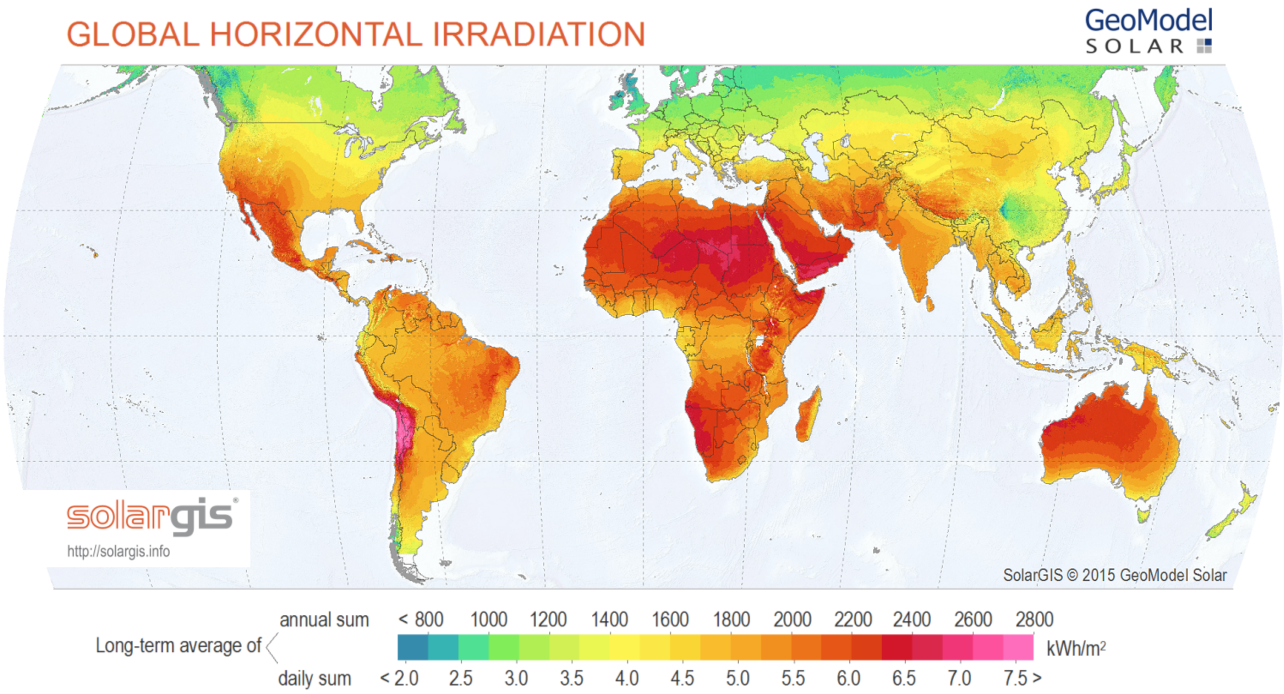

Figure 2.17: Annual global horizontal irradiation throughout the world [44].

2.4.2 Solar power profiles

The actual power output of a solar PV plant is, of course, influenced by a range of factors besides the available solar irradiance, including heat and conversion losses, panel and inverter specifications, panel temperature and maximum-power-point-tracking capabilities. Even for large solar PV installations that contain a multitude of system components, however, the local solar irradiance profile remains the strongest indicator of plant performance [41]. Since this study is exploratory rather than aimed at obtaining results for specific, real-world scenarios, it was therefore deemed unnecessary to include plant-specific parameters in the optimisation process. Consequently, each optimisation problem case considered in this project uses the solar irradiance profiles corresponding to the selected seasonal period and set of locations as the sole performance indicator for hypothetical solar PV power generation. In the context of each problem case, these irradiance profiles are therefore treated as a scalable equivalent of the power output profile for a solar PV installation of arbitrary size.[2] As a result, the power profile for a given distribution of generation capacity is represented by the aggregate (i.e. the sum) of the power profiles for each potential location scaled according to its per-unit capacity allocation.

In order to reflect the performance of potential solar PV plants realistically, GTI profiles of averaged hourly resolution are used to represent the power performance for the selected locations. It is a general guideline of solar PV system design that an elevation angle equal to the local latitude combined with an azimuth angle of zero (i.e. facing due north in the southern hemisphere) results in optimal annual energy collection for fixed-tilt PV panels [39]. The seasonal GTI profiles for each location were therefore obtained using these parameters, since

[2]For the sake of convenience, the term *power* will be used interchangeably with *irradiance* throughout the rest of the document, while *energy* will refer to *irradiation*.

they reflect realistic design configurations for hypothetical fixed-tilt solar PV installations. A combination of configurations that are somewhat eastern- or western-facing would likely be more complementary to the national demand profile, but the added complexity of optimising such an approach was deemed beyond the scope of this project. Naturally, the incident solar irradiance profiles associated with tracking PV panels would also be significantly different, but the dominant industry presence and lower costs associated with fixed-tilt PV plants [45] make them a more relevant focus point for this analysis.

2.5 Optimisation objectives

2.5.1 Overview

As stated in 2.1, the concept of grid support concerns both the extent of energy generation during peak demand periods, as well as the variability of the power profiles corresponding to solar PV generation. In view of these considerations, the following five optimisation objectives were identified for evaluation in the context of this study:

- Maximisation of the daily averaged energy
- Minimisation of the coefficient of variation (CV) for the daily averaged power profile
- Maximisation of the daily averaged energy weighted according to Eskom's Megaflex tariff structure
- Minimisation of the standard deviation (SD) for the cumulative daily energy
- Maximisation of the negative skewness of the probability distribution for the cumulative daily energy

Each of these optimisation objectives was derived to explore a significant aspect of the potential grid support applications of solar PV power generation, particularly with regard to the discussed seasonal and diurnal cycles. The first three objectives focus on the optimisation of energy generation for various seasonal and diurnal considerations, while the latter two objectives address the variability associated with solar power profiles on a seasonal basis. The various optimisation scenarios associated with each objective are implemented via their corresponding objective functions, which are presented in detail in Section 3.5.

2.5.2 Maximisation of the daily averaged energy

This optimisation objective aims to maximise the average availability of energy within a specified diurnal interval throughout a given seasonal period. In other words, the per-unit allocation of generation capacity is optimised to produce the highest cumulative energy value over a specified diurnal interval in the daily averaged aggregated power profile. In this context, the daily averaged profile refers to the 24-hour power profile produced by averaging the values for each hour over the full annual or seasonal duration of the original profile. The use of these averaged power profiles is in accordance with the long-term focus of this project, which is concerned with the improvement of aggregated solar PV performance on average rather than for any particular day.

With regard to the diurnal consideration, three specifications are evaluated for this optimisation objective, namely the full day scenario, the morning peak period and the evening peak period. All four seasons as well as the full year scenario are considered for both the full day and morning peak period specifications. As indicated in Section 2.2.2, however, the solar energy available during the evening peak period is limited and as such only the summer and full year scenarios are considered for this specifications.

2.5.3 Minimisation of the coefficient of variation for the daily averaged power profile

This optimisation objective aims to minimise the CV [p.u.] associated with a specified diurnal interval in the daily averaged aggregated power profile of a given seasonal period. The CV of a dataset, also known as the relative standard deviation (RSD), is defined as the ratio of the SD to the mean value and can be expressed per unit or as a percentage[3]. The formula for finding the CV of a dataset is as follows [46]:

$$\mathrm{CV} = \frac{\sigma_x}{\mu_x} \tag{2.2}$$

with

$$\sigma_{\boldsymbol{x}} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu_{\boldsymbol{x}})^2}{n-1}} \tag{2.3}$$

where n denotes the number of data points in the data vector $\boldsymbol{x}$,

$\mu_{\boldsymbol{x}}$ denotes the mean of $\boldsymbol{x}$,

and $\sigma_{\boldsymbol{x}}$ denotes the SD of $\boldsymbol{x}$.

Minimisation of this value therefore incorporates maximisation of the mean value for a set of values with simultaneous minimisation of the SD for the dataset, thereby implementing a multi-objective approach as a single minimisation problem. In theory, the aim of optimising the distribution of PV generation capacity according to this objective is to produce an aggregated daily averaged power profile characterised by a wide, flat shape. Compared to the solar power profiles typically associated with a single location, a power profile of this nature should have a lower peak power level in the middle of the day with more power available in the early morning

[3]For the sake of clarity, the term CV will be used to express per-unit values in the remainder of this document, while RSD will refer to percentage values.

and late afternoon hours. Since the maximisation and minimisation objectives are integrated rather than separately defined and weighted, however, there is no guarantee as to which element will be prioritised when converging to an optimal solution. The performance and usefulness of the described optimisation objective in this context is a significant aspect of this exploratory analysis.

Daily averaged profiles are once again utilised in order to reflect the average performance throughout a given seasonal period. The diurnal specifications considered for this approach consist of the full day scenario, for which all four seasons and the full year are evaluated, and the combined peak periods, for which only summer and the full year are evaluated.

2.5.4 Maximisation of the daily averaged energy weighted according to Eskom's Megaflex tariff structure

This optimisation objective also aims to maximise the average availability of energy throughout a given seasonal period, but incorporates Eskom's Megaflex tariff structure and its corresponding TOU schedule for weekdays as a weighting function. The energy profile associated with the aggregated daily averaged power profile of a distribution is therefore scaled according to the Megaflex structure to produce a prioritised cumulative energy value. Due to the nature of the TOU schedule, this approach includes an inherent diurnal cycle and as such is only evaluated for the full day scenario. In the seasonal context, all four seasons as well as the full year scenario are considered, with the weighting function adapted according to the Megaflex specifications for the high and low demand seasons.

2.5.5 Minimisation of the standard deviation for the cumulative daily energy

This optimisation objective simply aims to minimise the SD for the cumulative daily energy throughout a given seasonal period[4]. In other words, the per-unit allocation of generation capacity is optimised to produce the lowest variation (relative to the mean value) of the cumulative aggregated energy available for each day over the duration of a given seasonal interval. This should serve to increase the density and decrease the range of the mean-relative probability distribution histogram for the cumulative daily energy, as visualised is Figure 2.18. The value of this optimisation objective is apparent: lower day-to-day variability in the aggregated solar PV generation profile increases its reliability as a power source, while reducing the risks presented to the grid by large and/or frequent fluctuations of solar PV power output.

[4]In the context of this study, cumulative daily energy refers to the total energy available to a distribution over a 24-hour daily period. The cumulative daily energy throughout the duration of a seasonal period thus refers to the set of cumulative energy values for each day in the period.

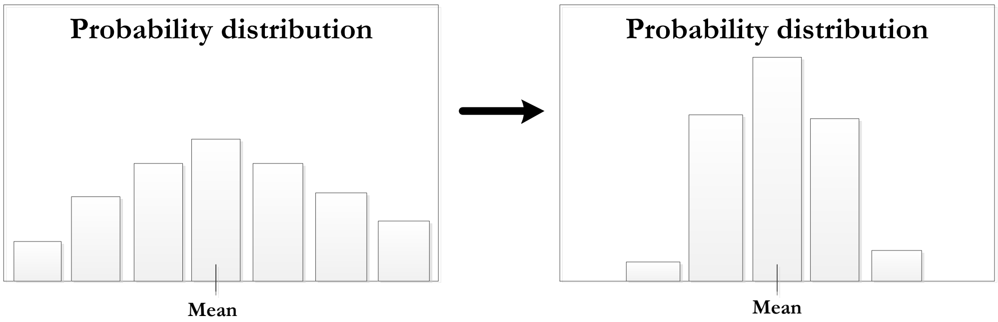

Figure 2.18: Visualisation for minimisation of the SD for the cumulative daily energy.

2.5.6 Maximisation of the negative skewness of the probability distribution for the cumulative daily energy

This optimisation objective aims to maximise the negative (i.e. right) skewness of the probability distribution for the cumulative daily energy throughout a given seasonal period. Essentially, this means the per-unit allocation of generation capacity is optimised so that the cumulative aggregated energy available on the majority of days throughout a seasonal period is as close as possible to the maximum cumulative energy delivered on any one day. When considering the probability distribution of the cumulative daily energy relative to the maximum cumulative energy value, this optimisation should serve to increase the negative (i.e. right) skewness of the histogram in the manner illustrated by Figure 2.19.

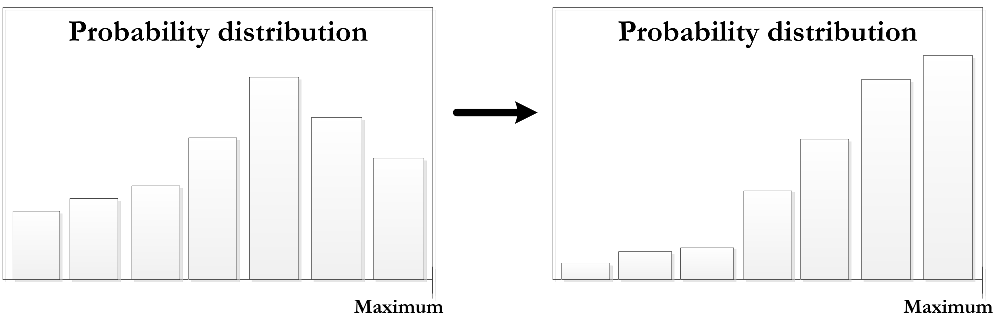

Figure 2.19: Visualisation for maximising the negative skewness of the probability distribution for the cumulative daily energy.

2.6 Optimisation algorithms

2.6.1 Overview

The varied and dissimilar nature of the selected optimisation objectives, combined with the discrete nature of the solar power profiles used to evaluate said objectives, serve to create a complex set of search problems. The nature of these search-spaces is multi-dimensional and discontinuous, with unknown topologies that may contain a myriad local minima. In this context, the implementation of classical derivative-based optimisation methods such as gradient descent or the Newton method are not feasible. Meanwhile, the application of exhaustive, derivative-free search algorithms for finding the global optimum of each search problem is also undesirable, since these methods can quickly become extremely expensive when increasing the size of multi-dimensional search spaces [47].

The preferred approach for this project is an optimisation algorithm that is directly applicable for any of the objective functions associated with the given optimisation objectives, as well as different objectives that may arise in future development. In addition, it should be relatively inexpensive in terms of both time requirements and computational resources, while remaining fairly robust with regard to performance over all problem cases. In view of these considerations, the field of metaheuristic optimisation algorithms was identified as the ideal source for suitable candidate algorithms.

2.6.2 Metaheuristics

This section discusses the subject of metaheuristics as presented in the book by Rothlauf [48], although similar information is available in a wide range of literature [49, 50]. Metaheuristics, also known as modern heuristics, is a subcategory in the field of heuristic optimisation methods. These methods were developed as an alternative to standard numerical optimisation methods for instances where the difficulty of solving a problem increases exponentially with the problem size. In many such cases, the resources and/or time required by traditional optimisation methods to find an optimal solution quickly become impractical as the problem size increases. Heuristic methods, in contrast to standard optimisation techniques, do not guarantee an exact solution, but usually have much lower resource requirements.

Three types of heuristic optimisation methods can be distiguished, namely: heuristics, approximation algorithms and metaheuristics. Heuristics are designed to use problem-specific (i.e. heuristic) information to solve optimisation problems quickly and deliver reasonably good solutions without guaranteeing optimality. Proper exploitation of problem knowledge in the design of the search method is usually necessary to ensure high-quality solutions. Heuristic optimisation is therefore less concerned with finding optimal solutions than developing optimum solution processes.

Heuristics can be implemented as either *construction* or *improvement* heuristics. The former constructs the elements of a complete solution through iterative construction steps, while the latter initialises a complete solution which is then iteratively improved via search operators. Approximation algorithms are simply heuristics that include a bound on the solution quality, i.e. construction or improvement heuristics that return an approximate solution of guaranteed quality.

Improvement heuristics that implement a problem-invariant search strategy are defined as metaheuristics. These methods offer the distinct advantage of being applicable to a fairly wide and diverse range of problems due to their problem-independant definition/design. Metaheuristics typically use alternating *intensification* and *diversification* steps throughout the search process, the first of which increases the focus on promising familiar areas in the search space, while the second explores new areas of the search space. The aim of intensification is to lead the search in the direction of solutions with higher fitness, while diversification aims to escape local optima.

Metaheuristics are characterised by the following design elements:

Representation — The representation describes how problem solutions are encoded in the context of the search space. It must be designed to accurately represent an optimal solution as well as enabling the application of variation operators to solutions.

Variation operator(s) — One or more variation operators, also known as search operators, are iteratively applied to represented solutions to generate new solutions. Variation operators can produce a new solution either by constructing one similar to an existing solution or by recombining properties of two or more existing solutions.

Fitness function — The fitness function is used to assess the quality of solutions, i.e. to distinguish between high- and low-quality solutions. The objective function of a search problem is often directly used as its metaheuristic fitness function, resulting in a solution fitness that is equal to its objective value.

Initial solution(s) — As metaheuristics are a type of improvement heuristics, one or more initial solutions are required to initiate the optimisation process. Initial solutions are often generated randomly, but high-quality initial solutions can also be created based on problem-specific knowledge or via an improvement heuristic.

Search strategy — The search strategy defines the operation and sequence of the intensification and diversification steps previously mentioned. These steps can either occur in an explicit progression or run in parallel during the iterative optimisation process.

Search strategies can be classified as either local search methods or recombination-based search methods. Local search methods, also known as direct search methods, aim to find an optimum point in the search space by iterating across neighbouring solutions, usually with the use of only one initial solution. These methods therefore rely primarily on intensification for their search strategy. A weakness of this approach is that the search can easily be led to a local rather than global optimum, especially if there are many local optima within the space. Consequently, various diversification strategies are incorporated by different metaheuristics to minimise the likelihood of this pitfall. Such strategies include the use of controlled diversification steps, modification of the representation, search operator or fitness function, and conducting repeated searches using different initial solutions.

In contrast to direct search methods, recombination-based search methods aim to converge a population of solutions to an optimal solution over a series of iterations. At each iteration the current population is used to create a new set of solutions through the implementation of a selection and subsequent recombination process. High-quality individuals are normally strongly favoured by the selection process to serve as parent solutions, thereby ensuring that favourable characteristics are retained within the population while undesirable characteristics are eventually eliminated. Once the pool of parent solutions is established, characteristics of selected individuals are recombined by recombination operators to produce a population of offspring solutions. The population is converged once all solutions possess the same characteristics, i.e. once the population is homogeneous. If the population converges to and consequently cannot escape a local optimum, it is referred to as premature convergence.

The main source of diversification in these search methods is the creation of a diverse initial population of solutions; thereafter, recombination operators iteratively enforce intensification by discarding characteristics that seem to produce low-quality solutions. Once again, various strategies can be implemented to increase diversification and the probability of avoiding premature convergence. Options include increasing the population size in order to increase the initial population diversity, limiting the level of intensification enforced by the selection process, or modifying the implementation of one or more of the design elements.

Due to the nature of the problems considered and the described analysis objectives, metaheuristics were deemed the most suitable optimisation approach for this study. Firstly, since the focus is on investigation rather than the development of exact solutions, the speed of heuristic methods is more desirable than the optimality of conventional methods. Furthermore, the problem-invariant nature of metaheuristic search techniques is preferable over both traditional heuristics and most classical optimisation methods since it facilitates the implementation of multiple objective functions within the same optimisation and simulation framework. With this approach, a range of possible objective functions can be evaluated using different metaheuristics without requiring individual design for each combination or adhering to a rigid problem structure.

The limitation of metaheuristics is of course that optimal or even reasonably good solutions are not guaranteed. Apart from the challenges presented by search spaces with many local optima, solution quality depends on proper selection and parameter adjustment of the search method in the context of the optimisation problem. Consequently, any given metaheuristic that is reasonably successful in solving a wide range of search problems, might deliver poor solutions for a different type of problem.

Although the optimisation algorithms in themselves are not the main subject of this study, evaluating the effectiveness of selected metaheuristics in solving the problem cases described in this chapter is a highly relevant analytical aspect. To this end, multiple methods utilising both types of search strategies were selected and applied to each problem. From among the recombination-based metaheuristics the popular genetic algorithm (GA) was selected, while pattern search optimisation was chosen as the candidate direct search method.

2.6.3 Genetic algorithm

Since its initial introduction by Fraser [51] in 1958 and further development by Bremermann [52] and Holland [53] in 1958 and 1975 respectively, the GA has been established as a powerful optimisation tool that can be applied to a diverse range of problems. Today the technique is widely used throughout various scientific fields and has been frequently and thoroughly described in literature [54, 55]. The rest of this section describes the principles and operation of simple GAs as presented in the book by Burke and Kendall [56].

Simple GAs incorporate the principles of genetics to simulate the processes of biological reproduction and natural selection in order to evolve good solutions for a given search problem. The algorithm creates a population of candidate solutions within the search space, after which the evolution of solutions is guided over a series of iterative steps using a problem-specific objective function as fitness measure to discriminate between good and bad solutions. Objective functions in this context are usually referred to as fitness functions. Candidate solutions take the form of strings of alphabets, which are encoded with the decision variables of the problem. The strings (i.e. solutions) are known as *chromosomes*, the alphabets on each string are referred to as *genes* and the values of genes are known as *alleles*.

The process to evolve good solutions consists of the following steps:

1. *Initialisation* The initial population of candidate solutions is generated across the search space, either randomly or based on problem-specific knowledge. Population size is an important consideration when using GAs as it can affect both the scalability and performance of the algorithm. A population that is too small might lead to premature convergence, resulting in substandard solution quality. On the other hand, overly large populations could result in unnecessary usage of computational time and resources.
2. *Evaluation* Once an initial or offspring population of candidate solutions are created, the relative fitness of each solution is evaluated using the specified objective function.
3. *Selection* A mating pool of individuals is selected from within the current population based on their fitness values. Numerous selection methods have been proposed to perform this step, all of which aim to allocate more slots to solutions with higher fitness values. By favouring good solutions over bad ones, the selection step incorporates the survival-of-the-fittest mechanism inherent to natural selection in biological evolution.

4. *Recombination* Segments of two or more parental solutions are combined to create new, potentially better offspring solutions that are not identical to any of their parents. This step can be executed either with a problem-specific strategy or using one of the many generic recombination (also known as crossover) operators already established. Appropriate design or selection of the recombination operator in the context of the search problem is essential to ensure satisfactory performance of the algorithm.
5. *Mutation* Once the recombination step has been completed, individuals are locally and randomly modified via mutation. Once again, there are a variety of mutation operators, all of which make changes to the genes of a candidate solution by performing a random walk in its vicinity. Although mutation is usually a secondary operator in GAs and occurs with low probability, it offers the advantage of increased diversity within the population and facilitates exploration throughout more of the search space.
6. *Replacement* The original population is replaced by the offspring population created via the selection, recombination and mutation steps. Depending on the replacement technique used, the old population is either replaced in its entirety or a specified number of individuals are replaced.
7. *Iteration* Steps 2 - 6 are repeated until a terminating condition is met.

2.6.4 Pattern search

The abstract definition of pattern search methods for solving non-linear unconstrained optimisation problems was first introduced by Torczon [57] in 1997, before being further expanded by Lewis and Torczon for bound constrained problems [58] and linearly constrained problems [59]. These pattern search methods, which became known as generalised pattern search (GPS) algorithms, explore a given search domain by means of a mesh structure in order to generate a series of iterates with non-increasing objective function values until convergence is achieved. A unified and convenient description for GPS algorithms is presented in the convergence analysis by Audet and Dennis Jr. [60]. Using this description, the GPS algorithm for a linearly constrained search problem with objective function $f(x)$ consists of the following steps:

1. *Initialisation* An initial point x_0 is defined so that $f(x_0)$ is finite, the initial mesh M_0 with mesh size parameter $\Delta_0 > 0$ and positive spanning directions set D_0 is defined on the search space, and the iteration counter k is set to zero.
2. *Search iteration* For each iteration k, an optional SEARCH and/or POLL step is performed to attempt to find a feasible, improved mesh point x_{k+1} on the mesh M_k, i.e. an iterate x_{k+1} that satisfies the search space constraints as well as the condition $f(x_{k+1}) < f(x_k)$.

2.1 SEARCH *step* The objective function f is evaluated at a finite number of trial points on the mesh M_k to find one that yields a lower value than the incumbent. Any user-defined search strategy may be applied for this step as long as a finite number of points are selected.

2.2 POLL *step* If an improved mesh point is not found during the SEARCH step, the POLL step invoked to evaluate f at a poll set of mesh points neighbouring the current iterate x_k in the directions of the set D_k. Once an improved mesh point is found, the polling process stops immediately. If the poll step fails to produce an improved mesh point, the incumbent solution is defined as a *mesh local optimiser*, i.e. its objective function value is less than or equal to that neighbouring mesh points.

3. *Parameter update* If the SEARCH or POLL steps produced an improved mesh point, the mesh size parameter is either kept the same or coarsened (i.e. $\Delta_{k+1} \geq \Delta_k$) so that upcoming SEARCH steps might explore more distant areas of the search domain. If x_k is a mesh local optimiser, the mesh size parameter is refined (i.e. $\Delta_{k+1} \leq \Delta_k$) so as to focus the search towards the region of the incumbent solution. After updating the mesh, the counter k is increased, x_{k+1} is set to x_k, and the next search iteration is initiated. Steps 2 - 3 are repeated until a terminating condition is met.

The key advantage of this method lies in the SEARCH step, since it accommodates the use of almost any heuristic or metaheuristic technique for performing random searches on the mesh. In view of this inherent flexiblility associated with pattern search algorithms, three different variations were selected for implementation and analysis with regard to the described optimisation objectives. The variations are specified as follows:

1. *No* SEARCH *step:* This variation is the simplest implementation of the pattern search method, with the optional SEARCH step omitted entirely.

2. SEARCH *step with generating set search (GSS):* The GSS method follows the same approach as the GPS algorithm for finding improved mesh points unless the current point is near a constraint boundary [61]. As such, it is often more efficient for finding optimal solutions for linearly constrained search problems and offers an advantage for the problems considered in this study.

3. SEARCH *step with GA:* This variation is essentially a hybrid of the GA and GPS methods and therefore offers a useful comparison to optimisation using the individual methods.

2.7 Summary of problem cases

When combined, the optimisation parameters discussed in this chapter produce a varied set of problem cases which all share a common design. Figure 2.20 presents the design parameters along with the logical progression that produces each distinct problem case.

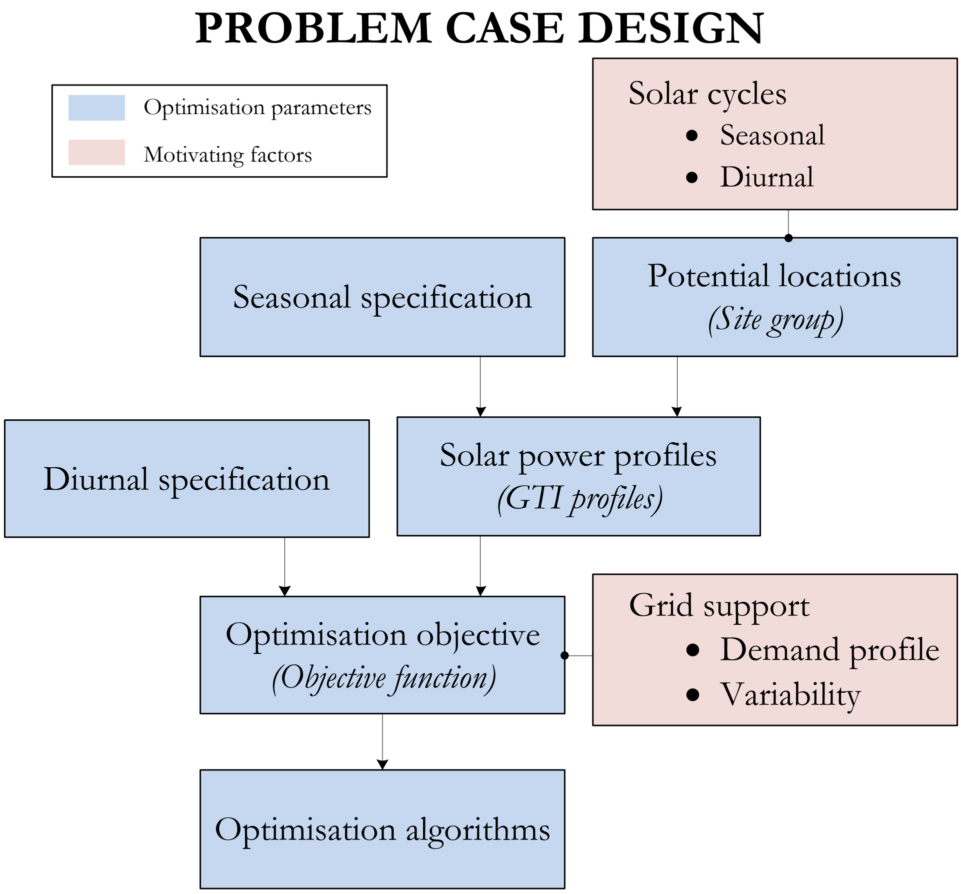

Figure 2.20: Design of problem cases in terms of optimisation parameters.

Table 2.7 provides a summary of the optimisation parameter sets that characterise the various problem cases identified for evaluation in this chapter. In addition to the parameters specified in the aforementioned table, each of the indicated problem cases was considered for all three site groups and was evaluated using all four of the described optimisation techniques. Due to the non-deterministic nature of the GA, a total of 50 evaluations were performed with each problem case for both pattern search variation 3 and GA optimisation.

Table 2.7: Summary of simulation parameters for problem cases.

Objective function	Diurnal period	Season				
		Autumn	***Winter***	***Spring***	***Summer***	***Full year***
1	Full day	✓	✓	✓	✓	✓
	Morning peak	✓	✓	✓	✓	✓
	Evening peak				✓	✓
2	Full day	✓	✓	✓	✓	✓
	Combined peak periods				✓	✓
3	Full day	✓	✓	✓	✓	✓
4	Full day	✓	✓	✓	✓	✓
5	Full day	✓	✓	✓	✓	✓

Chapter 3

Implementation of optimisation methodology

3.1 Overview

Having established the optimisation strategy and resulting problem cases for this exploratory study, a suitable set of solar power profiles was acquired and the necessary objective functions were defined for evaluating problem cases in the context of each optimisation objective. The software infrastructure required for implementation was subsequently constructed. As described in Section 1.4, the scope of this project includes the creation of a solar photovoltaic (PV) optimisation module based on the described optimisation parameters, which is to be integrated into an existing database-driven software application as part of an ongoing, overarching software project. The purpose of producing this module is not only to facilitate evaluation of the optimisation problem cases described in Chapter 2, but also to establish a set of practically applicable analysis tools as the foundation for future studies. In this regard, the software implementation of the optimisation methodology is a meaningful aspect of the project as a whole.

Figure 3.1 illustrates the interaction of the various components in the integrated software platform used for the implementation of the optimisation study. The user application interface consists of an established main application interface with the solar PV optimisation module incorporated as one of a number of independent plug-in modules created for various analysis purposes. The user application interface interacts with the relational database to retrieve and store information and data profiles associated with optimisation problem cases, while also facilitating the actual optimisation of said problem cases on an external simulation platform. The established database management functionality of the main application was extended to the solar PV module to facilitate database transactions, while communications with the external simulation platform is conducted solely within the environment of the solar PV module.

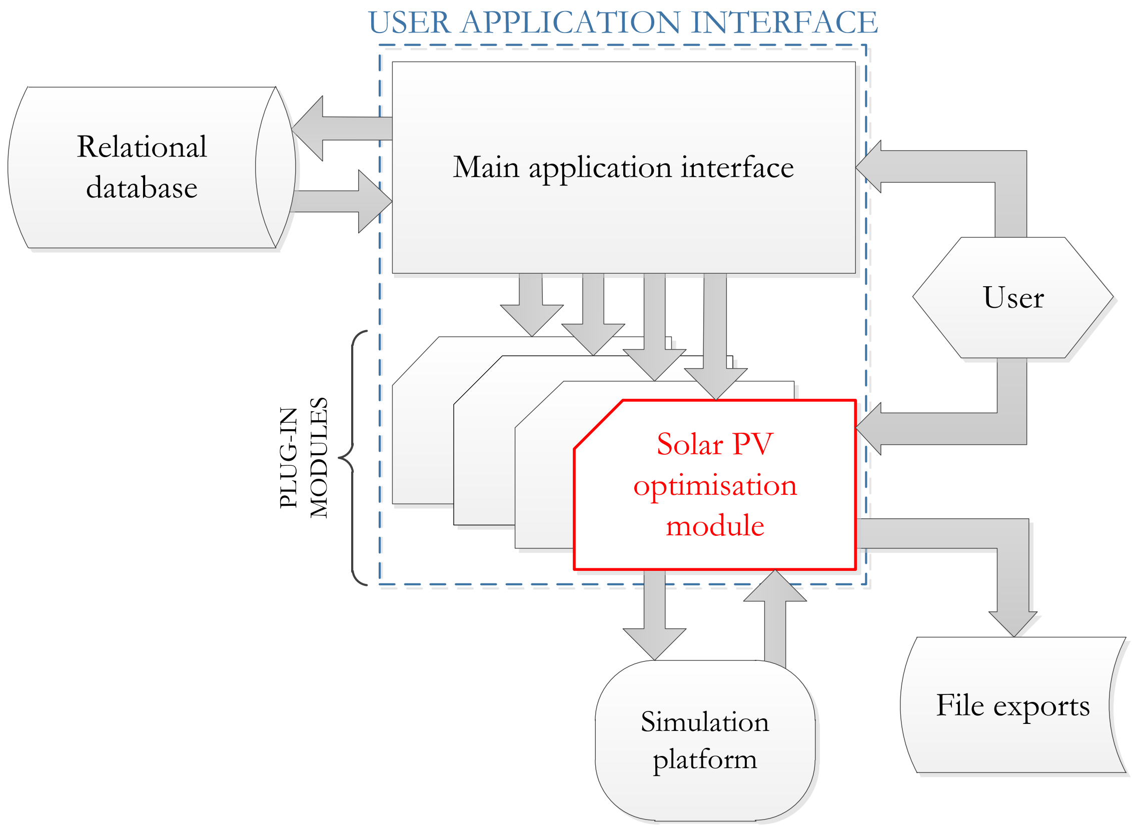

Figure 3.1: Diagram of integrated software platform.

With regard to the design of the relational database, the original relational database structure utilised by the main application interface was evaluated and partially adapted to better accommodate solar PV projects while remaining suitable for a wide range of diverse applications. Meanwhile, the necessary software functions for optimising the described problem cases in the context of the simulation platform were developed and incorporated into solar PV optimisation module. Apart from its role in the implementation of optimisation simulations, the module also includes analysis tools for evaluating the simulation results, as well as options for storing results in the database or exporting it to comma-separated values (CSV) files.

Figure 3.2 presents a diagram of the optimisation framework that applies to the evaluation of all optimisation problem cases. As indicated in the diagram, the framework consists of three steps, namely the input, simulation and output steps. In the context of the integrated software platform used to conduct each optimisation simulation, both the input and output steps are performed in the environment of the user application, while the simulation component is performed via the simulation platform.

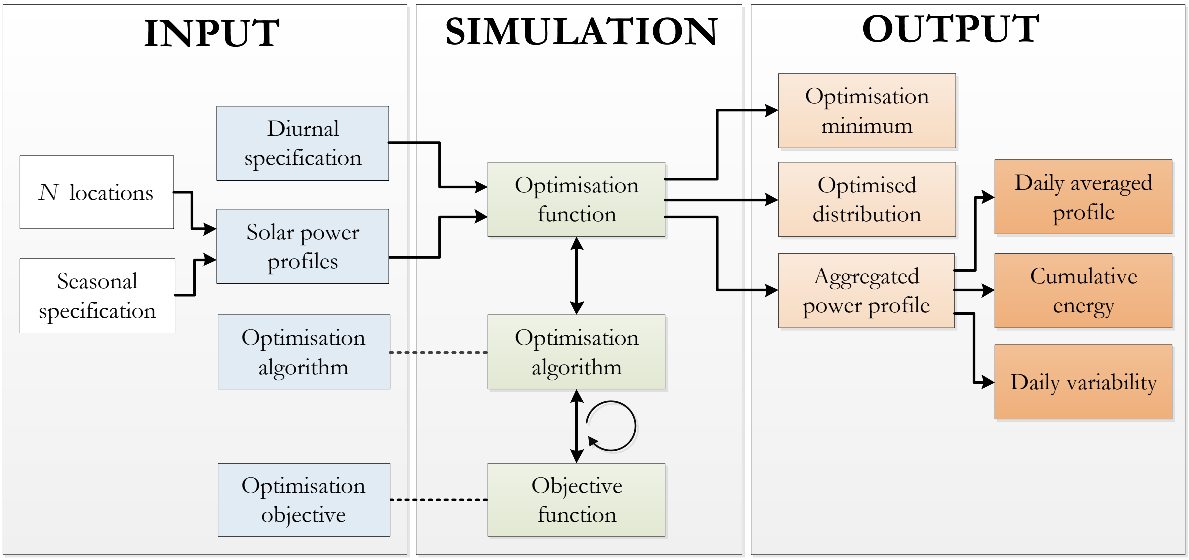

Figure 3.2: Diagram of optimisation framework.

For the input step, the user loads the solar power profiles for the relevant locations and seasonal period from the database and then selects the optimisation objective, diurnal specification and optimisation algorithm to be used in the simulation. When either the genetic algorithm (GA) or pattern variation 3 are selected as optimisation algorithm, the user also specifies the total number of evaluations to be performed. In the simulation step, all input parameters are passed to the optimisation software implemented for the simulation platform, which then attempts to find an optimal solution for the specified problem case. For the output step, the per-unit distribution produced by the optimisation simulation is returned to the user application interface, along with its corresponding aggregated solar power profile and objective function minimum. The user can then analyse these results via the solar PV optimisation module to produce secondary results such as the aggregated daily averaged profile.

3.2 User application interface

3.2.1 Overview

The software for the main application interface as well as the solar PV module was developed using the Embarcadero Delphi software development kit (SDK), which combines the Delphi programming language and Visual Component Library (VCL) framework within an Object Pascal integrated development environment (IDE). This development strategy was motivated by a number of factors, including the following [62]:

- Delphi offers extensive, robust functionality for graphical user interface (GUI) development coupled with very efficient compilers.

- Modular software development can easily be implemented in the form of different GUIs integrated into a root module, with the option of sharing common functional code where applicable.
- The object-oriented VCL framework offers a visual, highly intuitive approach for designing and implementing GUIs, which facilitates rapid development of user-friendly applications.
- Delphi offers strong database support and supplies various database drivers and components for the implementation of database-driven applications.
- Delphi software supports native cross-compilation as well as cross-platform development, which facilitates the ongoing expansion of the main application to include independent modules produced by different developers.
- External software developed with different programming languages can easily be incorporated into Delphi software through the use of dynamic-link library (DLL)s.

3.2.2 Main application interface

The key role of the main application interface within the integrated software platform is to perform user-driven database management and serve as the root application for all plug-in analysis modules. It offers the following database-related functionality:

- *Management of database connections:* The application can connect to a local or remote database that conforms to the expected database management system (DBMS) and structural configuration.
- *Navigation of the current database:* The application enables the user to explore connected databases in a step-wise fashion that follows the prescribed structural hierarchy presented in Section 3.3. At each level, the user can select any of the displayed data entries in order to retrieve and load associated entries in either the current level or one step down the hierarchy.
- *Retrieval of datasets:* After navigating to the lowest level of the database structure, i.e. the level at which large datasets or profiles are stored, the user can load a selection of these datasets to the application. The datasets can then be used for analysis purposes in a plug-in module or exported to files.
- *Alteration and creation of database entries:* The application allows the user to alter existing database entries and create new entries at each level of the structural hierarchy, while enforcing prescribed dependencies to maintain data integrity. The necessary information has to be entered manually by the user at each level except the last, for which data sets may be imported from CSV files.

3.2.3 Solar photovoltaic optimisation module

The features of the solar PV optimisation module can be characterised as follows:

- *Optimisation of the per-unit allocation of generation capacity for solarPV plant locations:* A user-defined problem case is evaluated via the external simulation platform to find the optimal distribution of per-unit solar PV generation capacity within a selection of locations. The relevant power profiles for the potential locations are selected and loaded from the database, while the user manually specifies the remaining optimisation parameters. The power profiles and optimisation specifications are passed to the simulation software, which performs the optimisation and returns the resulting per-unit distribution along with its aggregated power profile.
- *Analysis of the results produced for an optimisation problem case:* The aggregated power profile produced for an optimisation problem case is evaluated to determine one or more of the following:
  - The daily averaged power profile.
  - The total energy associated with the full day, morning peak period or evening peak period.
  - The standard deviation (SD) of the cumulative daily energy.
  - The probability distribution of the cumulative daily energy with regard to the mean value.
  - The probability distribution of the cumulative daily energy with regard to the maximum value.
- *Data storage of optimisation and analysis results:* The results returned by an optimisation simulation are either added to the database or exported in the form of a CSV file. Alternatively, the results produced for an analysis of the solution returned by an optimisation simulation are added tot the database or exported to a file.

3.3 Database development

3.3.1 Overview

For many software development purposes the use of a relational database offers a great deal of versatility compared to options such as a flat file system. In addition to offering the ideal structural platform for incorporating a range of data dependencies, the management of data is optimised via a DBMS and data integrity can be inherently guaranteed. The usefulness of a relational database for a particular application, however, depends a great deal on its structural design. This is of particular significance for the relational database used in the integrated software platform presented in Figure 3.1. Since the implemented database structure is in fact part of an ongoing software project that includes a variety of analytical features for a diverse range of applications, the design of the aforementioned structure had to be approached from this perspective. As such, the final database design is highly organisational, with a wide yet generic range of data fields that aim to accommodate all foreseeable data requirements.

3.3.2 The relational database model

By definition, an electronic database is a collection of information, preferably related, stored within a structure, preferably of an organised form [63]. Every database has both an intensional and an extensional component, with the former denoting its structural definition and the latter referring to the total data set stored within the database [64]. The intension of a database, often referred to as the database schema, is usually designed according to a specific database model. The first such models that emerged in the quest for efficient storage and management of electronic data was the hierarchical and network database models. The hierarchical database model is a simple inverted-tree structure with all links between entities taking the form of parent-child, one-to-many relationships. The network database model refines this approach with the addition of multiple-parent capability, which enables the mapping of many-to-many relationships [63].

The relational database model (RDM) was first introduced by Codd in 1970 [65] with the aim of facilitating the direct retrieval of data subsets from larger data sets [63]. In this respect, the RDM improves upon the limitations of the hierachical database model, which requires all data to be accessed via the root node of the structure. It also allows any part of the database topology to be linked to any other part, regardless of the data hierarchy. Over the ensuing years, the research of Codd and others served to establish the contemporary format of the RDM along with its general data access language, the Structured Query Language (SQL). Although various new database types, e.g. object-oriented databases and document-oriented databases, have been developed in recent years to better support specific applications, the RDM remains the most commonly used, straightforward and versatile approach for general purposes. As such, it was deemed the most appropriate database design stategy for the purposes of this project.

A relational database stores information in a set of tables, generally with each table representing a type of entity or object. The rows in a table are known as *records* or *tuples* and represent instances of that entity type, while the columns represent its various attributes. In order to preserve data integrity, each record in a table must be distinct and identifiable. To this end, each table contains an identifier column or set of columns called the primary key (PK), which consists of unique, not null values. When the data stored in two different tables is related, the primary key of one table appears in the related table as a foreign key (FK), thus linking associated records and enabling concurrent data retrieval across various tables [62].

An important characteristic of RDM design is the process of normalization, which eliminates duplication and data redundancy from a potential database schema via a series of normal forms. 1st Normal Form (1NP) removes repeating fields from a table by moving them to a new table that is linked to the truncated original in a parent-child, one-to-many relationship as depicted in Figure 3.3 [63]. In other words, attributes that exhibit consistently repeating values in the original table are relegated to a parent table, for which the primary key of each record can appear multiple times in the resulting child table as a foreign key.

Once the schema is in 1NP, 2nd Normal Form (2NP) is applied to move repeated values that are not functionally dependant on the primary key to new tables that are linked to the truncated original in a many-to-one relationship as shown in Figure 3.4 [63]. 2NP generally results in the creation of master-detail, many-to-one relationships between static and dynamic data, i.e. data that will stay much the same throughout the lifetime of the database versus data that will regularly be expanded or updated. Further normalization steps include 3rd Normal Form (3NP), Boyce-Codd Normal Form, 4th Normal Form, 5th Normal Form and Domain Key Normal Form [63], but these were deemed superfluous for the purposes of this project.

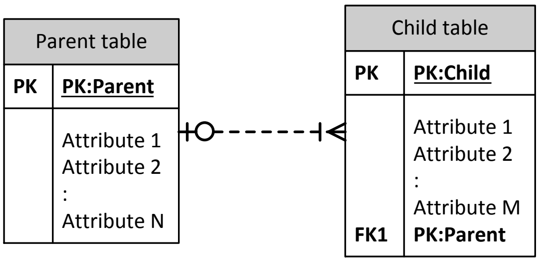

Figure 3.3: Result of 1NP.

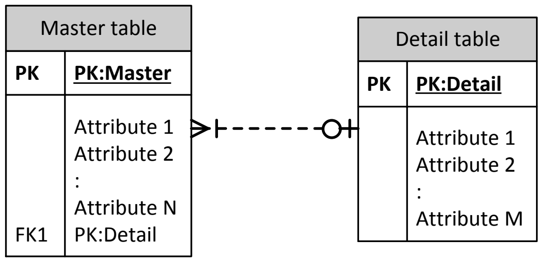

Figure 3.4: Result of 2NP.

3.3.3 Database management system

The implementation and management of any database is accomplished via a DBMS, which comprises the database engine as well as the software tools that controls interactions with the database [64]. A variety of SQL-based relational DBMSs are currently available, with some of the most popular being the following [66]:

- DB2
- Microsoft Access
- Microsoft SQL Server
- MySQL
- Oracle
- PostgreSQL
- SQLite

As one of the most widely-used relational DBMSs that is also available in various low-cost or free products, MySQL was identified as the most suitable DBMS candidate for the purposes of the ongoing software project. In addition to being low-cost and compatible with a wide range of operating systems and software, MySQL is user-friendly and can easily handle huge volumes of data, allowing for virtually unlimited data scaling [67]. The specific product selected for the database implementation performed in the context of the current project is the free, open-source software package WampServer, named for its four main components: the Windows operating system, the Apache HTTP Server, the MySQL relational DBMS and the PHP programming language. MySQL Workbench was subsequently selected as the software environment for database development, as it offers a convenient and versatile user interface.

3.3.4 Database architecture

In order to implement a detailed organisational structure that accommodates information and datasets of a diverse nature, the database was designed as a hierarchy of data components (i.e. entity types). These data components, which are classified as *Projects*, *Plants*, *Units*, *Profile Sets*, *Profiles* and *Profile Data*, produce a structure that is consistent with 1NP and therefore facilitates data handling and maintenance.

Each component consists of a main table supported by one or more auxiliary tables, which contain information pertaining to possible component *categories*, *tags* or, in the case of the *Profile* component, measurement *units*. Each component maintains a relationship to the components above and below it via link tables, also known as associative tables, while classifications from the supporting tables are included in the main component table via foreign keys. Figure 3.5 provides a visual representation of the overall component hierarchy.

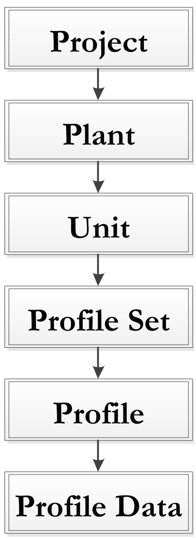

Figure 3.5: Top-to-bottom component hierarchy in the database structure.

3.3.4.1 Data components

Projects *Projects* are a mechanism for categorising data from one or more sources that serve a single user-defined purpose. All data stored in the database belong firstly to a *Project*, which contains one or more *Plants*. The *Project* component consists of the main *project* table presented in Table 3.1, as well as the supporting tables *projectcategory* and *projecttag*.

Table 3.1: Design of the *project* table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—
Comments	Text	—
CategoryID	Integer	Foreign
TagID	Integer	Foreign

Plants *Plants* refer to power plants, whether existing or potential, and are represented by the main *plant* table, shown in Table 3.2, along with the auxiliary tables *plantcategory* and *planttag*. All *Plants* belong to a *Project* via the link table *linkplant* and contains one or more units.

Table 3.2: Design of the *plant* table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—
Comments	Text	—
CategoryID	Integer	Foreign
TagID	Integer	Foreign
Location	Text	—
Latitude	Decimal	—
Longitude	Decimal	—

Units Power generation components of plants are classified as *Units* and are subordinated to a *Plant* via the link table *linkunit*. *Units* contain one or more *Profile Sets* of data and are fully described by the *unit* table, shown in Table 3.3, in conjunction with the *unitcategory* and *unittag* supporting tables.

Profile sets *Profile sets* consist of one or more *Profiles* from the same *Unit* that are grouped together for a user-defined purpose. *Profile sets* are assigned to *Units* via the link table *linkprofileset* and are represented by the main *profileset* table shown in Table 3.4 along with the auxiliary tables *profilesetcategory* and *profilesettag*.

Table 3.3: Design of the *unit* table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—
Comments	Text	—
CategoryID	Integer	Foreign
TagID	Integer	Foreign
Rating	Decimal	—

Table 3.4: Design of the *profileset* table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—
Comments	Text	—
CategoryID	Integer	Foreign
TagID	Integer	Foreign

Profiles A measured or simulated set of time series data over a specified interval is defined as a *Profile*, and belongs to a *Profile set* via the *linkprofile* link table. The *Profile* component consists of the main *profile* table shown in Table 3.5, the subsidiary tables *profilecategory* and *profiletag*, as well as the additional supporting table *profileunit*.

Table 3.5: Design of the *profile* table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—
Comments	Text	—
CategoryID	Integer	Foreign
TagID	Integer	Foreign
UnitID	Integer	Foreign

Profile data *Profile data* entries consist of single data points (i.e. a value and timestamp pair) that belong to a *Profile*. Since this is the bottom-level component in the database hierarchy, it will store the majority of data and handle the most SQL transactions. Consequently, *Profile Data* rows refer to their parent *Profile* via a foreign key in the main *profiledata* table rather than utilising a link table, as the intermediary table would significantly slow down queries and transactions from external sources when the database grows sufficiently large. The *Profile data* component is fully described by the *profiledata* table presented in Table 3.6 along with the *profiledatatag* auxiliary table. In this case, there is no need for a *category* table as the data is already categorised at *Profile* level.

Table 3.6: Design of the *profiledata* table structure.

Attribute	Data type	Key
ID	Integer	Primary
ProfileID	Integer	Foreign
Timestamp	Datetime	—
Value	Decimal	—
TagID	Integer	Foreign

3.3.4.2 Auxiliary tables

In accordance with 2NP, static aspects of the main data components are relegated to auxiliary tables that have a many-to-one relationship with the main component table. These tables act as classification systems for pertinent aspects of each component and as such take the form presented in Table 3.7. The value of *category* classifications is evident, as is that of the *profileunit* specification for the *Profile* component. Furthermore, *tag* classifications may prove useful for categorising or grouping components belonging to different parents according to specific user-defined traits.

Table 3.7: Design of the auxiliary table structure.

Attribute	Data type	Key
ID	Integer	Primary
Designation	Character string	—
Description	Text	—

3.3.4.3 Link tables

Link tables are generally utilised in 3NP for the mapping of many-to-many relationships to a single table. For this database structure, however, link tables featuring one-to-many relationships were deemed useful for separating data components in order to establish a more distinct hierarchy. Incorporating the link tables thus helps to preserve data integrity for large transactions extending through multiple levels of the database hierarchy, while also streamlining the main component tables. The table structure is straightforward, as shown in Table 3.8.

Table 3.8: Design of the link table structure.

Attribute	Data type	Key
ID	Integer	Primary
ParentID	Integer	Foreign
ChildID	Integer	Foreign

3.3.5 Database model

Figure 3.6 provides the entity relationship diagram (ERD) for the complete database structure, which is organised according to the described database components and clearly maps the column names, index keys and dependencies for each table.

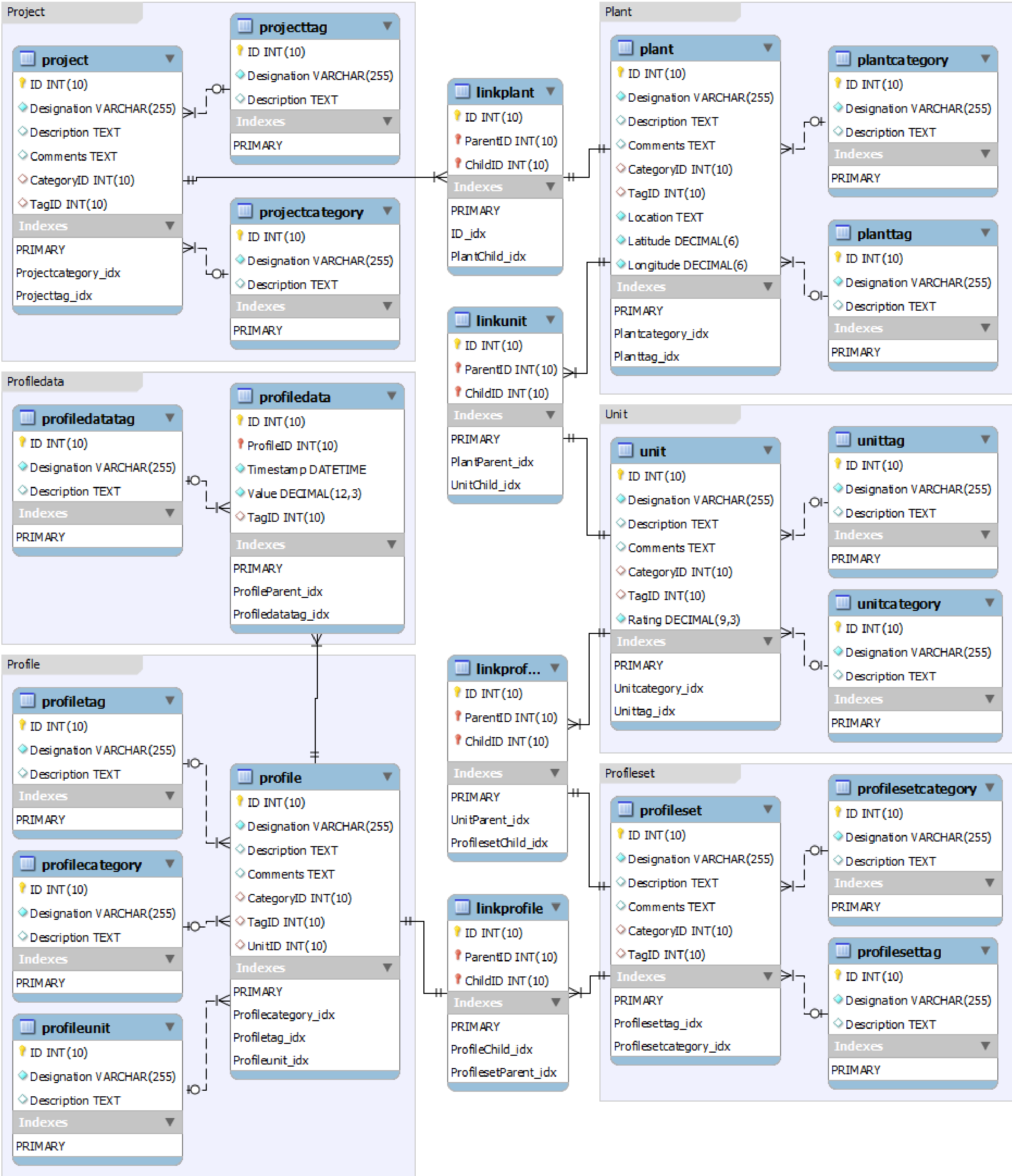

Figure 3.6: Entity relationship diagram (ERD) for the complete database structure.

3.4 Solar power profiles

Reliable long-term, high-resolution radiation data for South African locations is generally very difficult to obtain without incurring high costs [68]. However, since the goal of this optimisation study is not to obtain specific solutions but rather to evaluate the merits of the described approach, it is not essential to use actual measured values as long as the data displays realistic levels of seasonal and geographic variation. On this basis, all data used for the optimisation study was obtained from the PVsyst software package and its built-in meteorological data source.

The PVsyst application delivers solar radiation data as a set of synthetic hourly values for a typical meteorological year (TMY). To create these values for a selected location, the application first retrieves a set of monthly averages produced by long-term solar radiation measurements from the integrated METEONORM database. A random sequence of daily radiation values are generated from these average values using the model presented by Aguiar *et. al* [69], which incorporates a library of Markov matrices (also known as probability matrices) based on measured data. The autoregressive Gaussian model of Aguiar and Collares-Pereira [70] is then applied to the daily values to generate daily sequences of hourly values. The final yearly sequence of synthetic hourly values produced by these stochastic models consist of forward-filled averages, i.e. each data point represents the average radiation for the hour following its timestamp, with statistical properties that are analogous to measured solar radiation data [71].

For each location included in the site groups described in Section 2.3, the synthetic TMY dataset for the required Global Tilted Irradiance (GTI) (i.e. the GTI for a tilt angle equal to the local latitude and an azimuth angle of zero) was generated in PVsyst and exported to a CSV file. These files were processed to obtain the hourly GTI profiles for the individual seasons, each of which was stored in a separate CSV file. The complete set of hourly GTI profiles for the selected locations were then added to the database to be used in the relevant optimisation simulations.

With regard to the database structure, each site group was defined as a *Project*, while their various locations were defined as *Plants*. A single *Unit* with one *Profileset* was subsequently created for each *Plant*, with each *Profileset* containing five *Profiles* corresponding to the hourly GTI profiles obtained for the relevant location. Finally, the actual GTI profiles were imported to the database as *Profiledata* values belonging to the relevant *Profile*.

3.5 Objective functions

3.5.1 Overview

Each of the optimisation objectives defined in Section 2.5 require an equivalent objective function to evaluate the relative fitness of a given per-unit allocation during the optimisation process. The objective functions are defined as follows for each optimisation objective:

- *Objective function 1*: Maximisation of the daily averaged energy
- *Objective function 2*: Minimisation of the coefficient of variation (CV) for the daily averaged power profile
- *Objective function 3*: Maximisation of the daily averaged energy weighted according to Eskom's Megaflex tariff structure
- *Objective function 4*: Minimisation of the SD for the cumulative daily energy
- *Objective function 5*: Maximisation of the negative skewness of the probability distribution for the cumulative daily energy

Since the distribution of solar PV generation capacity is expressed on a per-unit basis, each optimisation problem is inherently characterised by boundary conditions specifying an allocation between the values of 0 and 1 for all locations, as well as the constraint that the aggregated allocation over all locations must have a value of 1. Furthermore, all optimisation objectives were implemented as minimisation problems, with the objective function fitness values simply taken as negative for maximisation objectives. The form of the search problems considered in this study can thus be defined as follows:

$$\begin{aligned} \underset{x}{\text{minimise}} \quad & f(x), \\ \text{subject to} \quad & 0 \leq x_n \leq 1, \quad n \in \{1, ..., N\}, \\ & \sum_{n=1}^{N} x_n = 1, \end{aligned} \tag{3.1}$$

where x_n denotes the per-unit allocation of generation capacity for each location n, and $f(x)$ denotes the objective function value.

Objective function 1 This function determines the aggregated daily averaged power profile for a given per-unit distribution of generation capacity and then calculates the cumulative energy [Wh/m^2] available over a specified diurnal period of that profile.

Objective function 2 This function determines the aggregated daily averaged power profile for a given per-unit distribution of generation capacity and then calculates the CV [p.u.] for a specified diurnal period.

Objective function 3 This objective function uses Eskom's Megaflex tariff structure and corresponding Time-of-Use (TOU) schedule for weekdays (previously shown in Table 2.2 and Table 2.1, respectively) as a weighting function to calculate a cumulative daily energy value for the aggregated daily averaged power profile of a given distribution. The weighting function is derived by representing the Megaflex tariffs for the relevant demand season as fractions of the maximum tariff, and then implementing these values as a vector of hourly weights [p.u.] in accordance with the TOU schedule. Each hourly value in the daily averaged energy profile [kWh/m^2] is then scaled according to the corresponding value in the weighting vector and summed to produce the cumulative daily energy. In the case of the full year analysis, the objective function calculates a weighted cumulative energy value for each of the four seasons using the relevant daily averaged profile, and then sums these values to produce the overall result. For this scenario, the weights of both the high and low demand seasons are expressed as a fraction of the maximum tariff for the high demand season.

Objective function 4 This objective function simply calculates the SD of the cumulative daily energy for a distribution throughout the relevant seasonal period. In contrast to the objective functions 1 to 3, the aggregated power profile is not resolved to a daily averaged profile but used to calculate the total energy delivered for each day. The SD of this dataset of cumulative daily energy values then serves as the objective function value.

Objective function 5 This objective function again finds the cumulative daily energy values for a given distribution and seasonal period, after which it normalises the dataset relative to its maximum value to produce a dataset of per unit values. The sum of these values then serves as the objective function value.

3.6 Simulation software

3.6.1 Simulation environment

All aspects of the simulation component in the optimisation framework shown in Figure 3.2 were implemented in the MATLAB simulation environment. The MATLAB platform was selected for this purpose based on its powerful computational capabilities, built-in simulation toolboxes and overall versatility. Novel software functions were developed for the general optimisation function as well as all five objective functions described in the previous section, while the GA and pattern search optimisation techniques were incorporated via the pre-defined functions available in MATLAB's Global Optimization Toolbox. The objective functions and optimisation algorithms are deployed within the general optimisation function, which was implemented as a DLL that can be accessed by the solar PV optimisation module.

3.6.2 Optimisation functions

3.6.2.1 GA function

The Matlab function declaration for finding a function minimum with the GA technique is given as [72]:

```
[x,fval] = ga(fitnessfcn,nvars,A,b,Aeq,beq,LB,UB,nonlcon,options)
```

with the input and output arguments defined as follows:

`x` — The best point in the search-space of `nvars` problem variables located by `ga` during its iterations. For the optimisation problems described in this chapter, this variable is a column vector of length `nvars` representing the optimal distribution of generation capacity found for a set of specified case parameters.

`fval` — The value returned by the fitness function for the point `x`, i.e. the optimisation minimum.

`fitnessfcn` — The handle to the fitness function defined for the problem, i.e. one of the objective functions defined in Section 3.5. Amongst its input arguments the fitness function must accept a row vector of length `nvars`, which in this case represents the fractional distribution of generation capacity, and it must return a scalar value, which is the described parameter calculated by each objective function.

`nvars` — A positive integer representing the number of variables in the problem, which translates to the number of sites for each problem case considered in this study.

`A, b` — The matrix A and vector b for linear inequality constraints of the form $A \cdot x \leq b$. Since the problems considered in this chapter contain no linear inequalities, these variables are specified as `[]` when calling the function.

`Aeq, beq` — The matrix A_{eq} and vector b_{eq} for linear equality constraints of the form $A_{eq} \cdot x = b_{eq}$. Since all problems in this study are subject to the constraint that the fractional distributions over all sites must add up to one, `Aeq` is a row vector of ones with length `nvars`, while `beq` is a scalar with the value 1.

`LB, UB` — Vectors of length `nvars` representing the lower and upper bounds for the problem variables. Since the fractional distribution for each site must always be between zero and one, the values of `LB` are set to 0 while the values of `UB` are set to 1.

`nonlcon` — The handle to a function that describes all non-linear problem constraints. Since this study contains no non-linear constraints, this variable is specified as `[]` when calling the function.

`options` — A structure of optimisation options defined with the `gaoptimset` function.

The `options` argument allows for the specification of a wide range of simulation parameters, most of which are either not pertinent or set to default specifications that an initial analysis deemed sufficient for the purposes of this study. Only the `MaxGenerations` parameter, denoting the maximum number of generations allowed before terminating the simulation, was increased from the default of 100·`nvars` to 100000·`nvars`.

3.6.2.2 Pattern search function

The function declaration for finding a function minimum with the pattern search method is given as [72]:

```
[x,fval] = patternsearch(fun,x0,A,b,Aeq,beq,LB,UB,nonlcon,options)
```

with `fun` denoting the handle to the relevant fitness function and `x0` denoting a vector specifying the initial point (i.e. initial distribution) for the search problem. As previously stated, the initial distribution for each problem case is defined as uniform with a value of $1/N$ for each location, where N represents the number of potential locations considered in the problem. The remainder of the input and output arguments are equivalent to their counterparts as defined for the GA function declaration.

Once again, the `options` argument allows for the specification of a wide range of simulation parameters, many of which are not pertinent or set to default specifications that an initial analysis deemed sufficient for the purposes of this study. The parameters that were adjusted include `InitialMeshSize`, which was decreased from 1 to 0.2, `MaxIterations`, which was increased from 100·`nvars` to 100000·`nvars`, and `MaxFunctionEvaluations`, which was increased from 2000·`nvars` to 100000·`nvars`. The search function parameter was also set to generalised pattern search (GPS), generating set search (GSS) and GA for pattern search variations 1, 2 and 3 respectively.

3.6.2.3 General optimisation function

The function declaration for the general optimisation function used to analyse each problem case is given by:

```
[dists, mins, cumprof] = simulation(nloc, power, ts, ...
                                    tou, season, objfun, optm, iter)
```

with the input and output arguments defined as follows:

`dists` An array with `nloc` rows and `iter` columns representing the optimal distributions found for each optimisation simulation performed.

`mins` A vector of length `iter` representing the optimisation minima associated with each solution in `dists`.

`cumprof`	An array with `iter` rows representing the cumulative seasonal power profile corresponding to each solution.
`nloc`	An integer denoting the number of potential locations considered.
`power`	Generally an array with `nloc` rows representing the seasonal power profiles for each potential location. When using objective function 1 for the full year scenario, this parameter becomes a cell array containing a set of seasonal `nloc` power profiles for each season.
`ts`	The timestamps corresponding to the values in `power` and represented by datetime arrays following the same structure as `power`.
`tou`	An array representing the start and end times of the optimisation TOU period(s) in term of hours on a scale of 0 to 24.
`season`	An iteger representing the seasonal period for which the optimisation is performed, with 1 denoting autumn, 2 denoting winter, 3 denoting spring, 4 denoting summer and 5 denoting the full year.
`iter`	An integer denoting the number of optimisation simulations to be performed for the problem case in question. The parameter defaults to 1 for pattern search variations 1 and 2.
`objfun`	An integer representing the objective function to be used for optimisation. Objective functions are identified according to their description number detailed in Section 3.5.

Chapter 4

Results and discussion

4.1 Overview

This chapter presents a detailed analysis of the results for all problem cases analysed during the implementation of the optimisation study. Discussions are presented separately with regard to performance evaluation of the optimisation algorithms and problem case results, with only the best solution found for each problem case considered in the latter analysis. In order to simplify the analysis process, the problem cases are grouped in terms of the optimisation objectives and, where applicable, further organised according to the diurnal specification.

4.2 Analysis of optimisation techniques

4.2.1 Overview

This section compares the performance of the selected optimisation techniques in solving the problem cases for each objective function. In each case the minima for pattern search variations 1 and 2 are considered along with the best and worst minima found for pattern search variation 3 as well as the genetic algorithm (GA). The results are expressed as the percentage deviation from the best overall minimum found by the techniques for each problem case, with a value of zero for each method that equalled the best minimum.

Apart from solution quality, speed is also a prevalent factor in assessing the effectiveness of the selected metaheuristic methods. Though the priority for an optimisation problem is to approximate the global minimum of the objective function, very small improvements to the solution may not be desirable if it greatly extends the simulation runtime. The duration of simulations for pattern search variations 1 and 2 proved to be negligible, and as such are omitted from the discussion, but the simulation times and multiple evaluations associated with pattern search variation 3 and the GA can result in a much more significant time consideration. To this end, the average duration of evaluations performed by pattern search variation 3 as well as the GA is also discussed for each objective function.

4.2.2 Objective function 1: Maximisation of the daily averaged energy

Table 4.1 presents the results for problem cases using objective function 1. For the full day scenario the overall best minimum for each problem case is the best minimum found by pattern search variation 3 and/or the best GA minimum. Overall, pattern search variation 3 performed slightly worse than the GA, with deviation for every winter case and fewer overall best minima. The worst minima for both these methods show some variation, with the highest deviation occurring for winter in site groups 1 and 3. The minima for pattern search variation 1 and 2 also show varying degrees of deviation relative to the overall best minima, again with particularly high instances for winter in site groups 1 and 3. Pattern search variation 2 generally performed better for site groups 1 and 2, while pattern search variation 1 shows the best solution quality for site group 3, with the exception of the winter case.

For these problem cases, the average simulation runtime for pattern search variation 3 was found to be up to 1 second for site groups 1 and 2 and up to 3.1 seconds for site group 3. Meanwhile the average duration of the GA simulations was up to 3.9 seconds for site groups 1 and 2 and up to 70.4 seconds for site group 3. Therefore, although the GA performed slightly better in terms of solution quality, pattern search variation 3 significantly outperformed it with regard to speed, particularly for site group 3.

For the morning peak period the results are more varied: all four methods produced multiple overall best minima, with the GA finding the most. Pattern search variation 2 shows the best performance for site group 1, but deviates greatly from the overall best minima for several problem cases in site groups 2 and 3, most notably for the winter cases. The GA consistently produced the best overall minima for site group 2 and also shows the best general performance for site group 3, although it deviates notably from the overall minimum for the winter case. The results for pattern search variation 1, as well as the worst minima for both pattern search variation 3 and the GA, show significant variation for all three site groups. For the latter two methods, this variation indicates that a significant number of evaluations are necessary to guarantee the current solution quality.

The simulations for pattern search variation 3 averaged a runtime of up to 1 second for site groups 1 and 2, and up to 3.9 seconds for site group 3. The GA simulations had an average duration of up to 6.5 seconds for site groups 1 and 2, and up to 96.9 seconds for site group 3. When considering the results for all site groups in terms of both solution quality and speed, no one method stands out as the best overall performer. Pattern search variation 2 produced the best combined set of minima for site group 1 with negligible runtime, while the GA produced the best overall minima for site group 2 with relatively short simulation times. For site group 3, the significantly lower runtime of pattern search variation 3 may render it preferable to the GA, even though the GA performed slightly better in terms of solution quality.

For the evening peak period, the GA produced the overall best minima for all problem cases, with pattern variation 3 equalling these minima for all but one case. Pattern search variations

1 and 2 produced varying results, with variation 1 showing significant deviation for site group 1 and 2 but not for site group 3, while the variation 3 minima greatly deviate for site group 2 and 3 but not for site group 1. The worst minima for pattern search variation 3 and the GA show significant variation only for site group 2, with deviations that are substantial but still better than those produced by pattern search variations 1 and 2.

Table 4.1: Percentage deviation of optimisation minima relative to the best minimum found for problem cases using objective function 1.

Problem case description			**Pattern search variation 1**	**Pattern search variation 2**	**Pattern search variation 3**		**GA**	
Diurnal period	***Site group***	***Seasonal period***			***Best***	***Worst***	***Best***	***Worst***
Full day	1	Autumn	2.439	0.854	0.043	0.865	0.000	0.875
		Winter	5.591	8.059	0.301	4.677	0.000	2.742
		Spring	3.476	0.005	0.000	0.037	0.000	0.026
		Summer	3.361	0.006	0.000	0.011	0.000	0.024
		Full year	1.904	0.230	0.000	0.232	0.030	0.253
	2	Autumn	1.638	0.002	0.000	0.031	0.000	0.025
		Winter	1.819	0.511	0.092	0.551	0.000	0.522
		Spring	0.198	0.344	0.000	0.351	0.026	0.368
		Summer	1.119	1.906	0.062	1.917	0.000	1.917
		Full year	0.462	0.000	0.000	0.022	0.000	0.026
	3	Autumn	0.533	1.326	0.000	0.538	0.529	0.547
		Winter	7.857	6.623	0.253	3.696	0.000	3.830
		Spring	0.005	1.736	0.000	0.014	0.000	0.014
		Summer	0.006	1.679	0.000	0.011	0.000	0.015
		Full year	0.217	1.052	0.212	0.224	0.000	0.228
Morning peak	1	Autumn	3.832	0.000	0.056	4.942	0.132	3.108
		Winter	2.292	0.024	0.438	7.257	0.000	4.998
		Spring	2.625	0.003	0.000	4.029	0.056	3.534
		Summer	2.855	0.000	0.006	4.999	0.068	3.674
		Full year	3.650	0.000	0.090	4.107	0.041	4.676
	2	Autumn	5.864	9.815	0.560	9.806	0.000	6.323
		Winter	11.496	19.204	0.390	10.998	0.000	6.564
		Spring	1.892	0.003	0.000	0.029	0.000	0.025
		Summer	1.315	0.003	0.000	0.249	0.000	0.015
		Full year	3.914	6.556	0.291	6.548	0.000	3.999
	3	Autumn	8.963	8.963	0.967	10.292	0.000	8.036
		Winter	18.708	18.708	0.000	13.492	1.296	11.031
		Spring	0.000	0.000	0.345	3.109	0.302	4.755
		Summer	0.000	0.000	0.119	2.252	0.080	1.974
		Full year	5.681	5.681	0.617	6.295	0.000	3.753
Evening peak	1	Summer	18.790	0.030	0.000	0.029	0.000	0.031
		Full year	0.013	0.019	0.000	0.026	0.000	0.024
	2	Summer	15.933	23.163	0.000	11.196	0.000	12.919
		Full year	16.959	27.062	0.479	13.017	0.000	8.117
	3	Summer	0.000	9.371	0.000	0.029	0.000	0.037
		Full year	0.000	13.755	0.000	0.021	0.000	0.019

The average simulation runtime for pattern search variation 3 was found to be up to 1 second for site groups 1 and 2, and up to 2.5 seconds for site group 3. Meanwhile the average simulation runtime for the GA was up to 4.7 seconds for site groups 1 and 2 and up to 53.7 seconds for site group 3. Given that pattern search variation 3 closely matched the performance of the GA in terms of accuracy, its obvious advantage in terms of simulation duration indicates that overall it is the more desirable optimisation technique for these problem cases.

4.2.3 Objective function 2: Minimisation of the coefficient of variation (CV) for the daily averaged power profile

Table 4.2 presents the results for problem cases using objective function 2. For the full day as well as the combined peak periods, pattern search variation 3 produced the overall best minima for the majority of problem cases with negligible deviation for the remaining cases. The GA performed almost as well, with its best minima either matching the overall best minima or showing negligible to very small deviations. Pattern search variations 1 and 2 both performed slightly worse than the GA by producing more significant and frequent deviations.

Table 4.2: Percentage deviation of optimisation minima relative to the best minimum found for problem cases using objective function 2.

Problem case description			Pattern search variation 1	Pattern search variation 2	Pattern search variation 3		GA	
Diurnal period	*Site group*	*Seasonal period*			*Best*	*Worst*	*Best*	*Worst*
Full day	1	Autumn	0.152	0.097	0.005	0.107	0.000	0.102
		Winter	0.047	0.237	0.000	0.745	0.002	0.908
		Spring	2.524	0.000	0.000	0.000	0.000	1.671
		Summer	0.063	0.085	0.000	0.101	0.002	0.738
		Full year	0.068	0.479	0.000	0.476	0.004	0.475
	2	Autumn	0.544	1.094	0.008	0.483	0.000	0.322
		Winter	0.229	0.000	0.000	0.561	0.011	1.060
		Spring	0.071	0.171	0.003	0.171	0.000	0.171
		Summer	0.061	0.000	0.011	0.069	0.014	0.306
		Full year	0.048	0.104	0.000	0.121	0.002	0.160
	3	Autumn	1.463	1.463	0.000	0.871	0.173	1.031
		Winter	2.285	2.162	0.000	1.933	0.180	1.833
		Spring	0.000	0.937	0.000	0.000	0.000	0.000
		Summer	0.003	0.000	0.015	0.504	0.103	0.787
		Full year	0.406	0.160	0.000	0.412	0.075	0.495
Combined peak periods	1	Summer	0.714	0.165	0.000	0.166	0.001	0.165
		Full year	0.315	2.213	0.000	2.106	0.001	1.191
	2	Summer	0.000	0.000	0.000	0.000	0.019	0.817
		Full year	0.092	0.000	0.003	0.224	0.011	0.918
	3	Summer	0.159	0.046	0.000	0.163	0.038	0.164
		Full year	1.982	1.007	0.000	1.927	0.084	1.510

With regard to the simulation runtime for these problem cases, pattern search variation 3 averaged a duration of up to 1.1 seconds for site groups 1 and 2, and up to 3.7 seconds for site group 3. The GA simulations averaged a runtime of up to 5.3 seconds for site groups 1 and 2, and up to 64.8 seconds for site group 3. Considering its solution quality as well as its average simulation times, pattern search variation 3 appears to offer the best overall performance for this objective function.

4.2.4 Objective function 3: Maximisation of the daily averaged energy weighted according to Eskom's Megaflex tariff structure

Table 4.3 presents the results for problem cases using objective function 3. For each problem case, the best overall minimum is either the best minimum found by pattern search variation 3 and/or the best GA minimum. The two optimisation methods show very similar general solution quality with regard to their best minima, while both also show notable variation in their worst minima, especially for the winter period. Pattern variations 1 and 2 both show a similar level of deviation, though not necessarily for the same problem cases.

Table 4.3: Percentage deviation of optimisation minima relative to the best minimum found for problem cases using objective function 3.

Problem case description		Pattern search variation 1	Pattern search variation 2	Pattern search variation 3		GA	
Site group	***Seasonal period***			***Best***	***Worst***	***Best***	***Worst***
1	Autumn	2.674	2.318	0.035	2.319	0.000	2.329
	Winter	3.629	0.007	0.000	3.360	0.076	3.778
	Spring	3.097	0.003	0.000	0.017	0.000	0.021
	Summer	2.555	0.005	0.000	0.023	0.000	0.023
	Full year	2.280	3.116	0.000	3.105	0.081	2.396
2	Autumn	1.824	0.004	0.000	0.016	0.000	0.032
	Winter	4.038	6.778	0.227	5.235	0.000	4.690
	Spring	0.036	0.000	0.000	0.014	0.000	0.024
	Summer	0.914	1.542	0.000	1.542	0.036	1.558
	Full year	1.022	0.001	0.000	0.015	0.000	0.015
3	Autumn	1.831	2.009	0.006	1.835	0.000	1.842
	Winter	6.255	6.255	0.000	6.569	0.075	5.428
	Spring	0.003	1.546	0.000	0.006	0.000	0.018
	Summer	0.005	1.276	0.000	0.010	0.003	0.013
	Full year	2.861	2.443	0.116	2.871	0.000	2.862

The average simulation runtime for pattern search variation 3 was found to be up to 6 seconds for site group 1 and 2, and up to 19.5 seconds for site group 3. The execution time for the GA simulations averaged at up to 38.6 seconds for site groups 1 and 2, and up to 327 seconds for site group 3. Given their similar performance in terms of solution quality, the significantly shorter simulation durations exhibited by pattern search variation 3 once again makes it a more advantageous option than the GA. It also offers consistently better solution quality compared to pattern search variations 1 and 2, provided that a sufficient number of evaluations are performed.

4.2.5 Objective function 4: Minimisation of the standard deviation (SD) for the cumulative daily energy

Table 4.4 presents the results for problem cases using objective function 4. For this objective function, all three pattern search variations consistently delivered the same minima, with the one exception of a small variation in the winter period for site group 3. For the GA results, the best minima show negligible or no deviation from those found by the pattern search variations, while the worst minima show more significant variation. With regard to the time consideration, the average simulation runtime for pattern search variation 3 was up to 1.4 seconds for site groups 1 and 2, and up to 4.8 seconds for site group 3. For the GA simulations, the average duration was up to 3.5 seconds for site groups 1 and 2, and up to 88.5 seconds for site group 3.

Table 4.4: Percentage deviation of optimisation minima relative to the best minimum found for problem cases using objective function 4.

Problem case description		Pattern search variation 1	Pattern search variation 2	Pattern search variation 3		GA	
Site group	***Seasonal period***			***Best***	***Worst***	***Best***	***Worst***
1	Autumn	0.000	0.000	0.000	0.000	0.000	0.214
	Winter	0.000	0.000	0.000	0.000	0.001	2.231
	Spring	0.000	0.000	0.000	0.000	0.001	0.268
	Summer	0.000	0.000	0.000	0.000	0.001	0.559
	Full year	0.000	0.000	0.000	0.000	0.000	0.141
2	Autumn	0.000	0.000	0.000	0.000	0.001	0.366
	Winter	0.000	0.000	0.000	0.000	0.001	0.262
	Spring	0.000	0.000	0.000	0.000	0.001	0.117
	Summer	0.000	0.000	0.000	0.000	0.002	0.159
	Full year	0.000	0.000	0.000	0.000	0.002	0.167
3	Autumn	0.000	0.000	0.000	0.000	0.002	2.121
	Winter	0.136	0.136	0.000	0.114	0.009	2.832
	Spring	0.000	0.000	0.000	0.000	0.003	0.561
	Summer	0.000	0.000	0.000	0.000	0.009	2.241
	Full year	0.000	0.000	0.000	0.000	0.004	1.246

In view of these results, all three variations of the pattern search method offer a clear advantage over the GA for this particular objective function, in terms of both solution quality and especially speed. With the exception of the slight variation for winter for site group 3, the results indicate that for each problem case all evaluations performed with pattern search variation

3 produced the same minimum. Thus it appears that the three pattern search variations yield the same solution for each of these cases, with pattern search variation 3 requiring only one evaluation to find this optimal solution. Given the fact that there is an exception to this trend, however, it cannot be confidently predicted that the pattern search methods would maintain this consistency and solution quality with different input data and/or simulation parameters.

4.2.6 Objective function 5: Maximisation of negative skewness of the probability distribution for the cumulative daily energy

Table 4.5 presents the results for problem cases using objective function 5. Once again, the overall best minimum produced for each problem case is either the best minimum found by pattern search variation 3 and/or the best GA minimum. On the whole, the GA performed worse than pattern search variation 3 since it produced fewer overall minima and shows greater deviation from the overall minima, especially for site group 3. The worst minima for both pattern search variation 3 and the GA show notable variation, in many cases producing worse solutions than one or both of pattern search variations 1 and 2.

Table 4.5: Percentage deviation of optimisation minima relative to the best minimum found for problem cases using objective function 5.

Problem case description		Pattern search variation 1	Pattern search variation 2	Pattern search variation 3		GA	
Site group	*Seasonal period*			*Best*	*Worst*	*Best*	*Worst*
1	Autumn	2.321	1.527	0.000	1.625	0.080	1.491
	Winter	0.324	0.744	0.043	0.635	0.000	0.457
	Spring	0.041	0.087	0.078	0.673	0.000	1.612
	Summer	0.987	0.001	0.000	0.006	0.000	0.006
	Full year	0.258	0.250	0.000	0.636	0.081	1.513
2	Autumn	0.382	0.382	0.000	0.434	0.023	1.581
	Winter	0.096	0.108	0.000	0.405	0.029	0.535
	Spring	0.800	0.802	0.014	1.118	0.000	0.601
	Summer	0.127	0.250	0.023	0.794	0.000	0.920
	Full year	0.075	0.075	0.000	0.473	0.013	0.402
3	Autumn	0.274	0.576	0.000	1.589	0.181	1.567
	Winter	0.482	0.860	0.000	1.243	0.615	1.880
	Spring	0.148	0.142	0.000	1.531	0.476	2.883
	Summer	0.057	0.251	0.000	0.226	0.222	0.955
	Full year	0.285	0.674	0.000	1.358	0.190	2.415

The average simulation runtime for pattern search variation 3 was found to be up to 2.1 seconds for site group 1 and 2, and up to 9 seconds for site group 3. The execution time for the GA simulations averaged at up to 6.7 seconds for site groups 1 and 2, and up to 195 seconds for site group 3. Pattern search variation 3 therefore once again outperformed the

GA in both speed and overall solution quality, most significantly for site group 3. Although it also consistently outperformed pattern search variations 1 and 2 in terms of solution quality, the variation of the worst minima again indicates that a sufficient number of evaluations are necessary to guarantee solution quality.

4.2.7 Summary

As expected, results for the selected optimisation techniques reveal that their performance in optimising all the selected problem cases varies somewhat, both compared to each other and for different optimisation functions and simulation parameters. Pattern search variations 1 and 2 proved inconsistent for all objective functions except objective function 4, with pattern search variation 1 exhibiting the worst overall solution quality. The minima for both these methods frequently show notable deviations from the overall best minima, especially for objective functions 1 and 3, which indicates that the search algorithms struggle to escape local minima in the search space of the problem cases considered.

Pattern search variation 3 shows the best overall solution quality for all five objective functions, with the most overall best minima and only small to negligible deviations. The GA delivered generally similar results, which are slightly better for certain sets of problem cases but worse overall. Although only the best minima were considered in evaluating the performance of both pattern search variation 3 and the GA, their worst minima offer some insight as to the consistency of the solutions and the difficulty of optimising for particular objective functions. Frequent variation, especially when reflected in the minima of pattern search variations 1 and 2, may indicate the presence of prominent and/or numerous local minima in the search space associated with specific data sets and objective functions.

The worst minima for many problem cases using objective function 1 shows significant deviation from the overall best minimum, with lesser and less frequent deviations for objective functions 2, 3 and 5. The greatest variation by far occurs for selected evening peak and winter morning peak problem cases, which can probably be ascribed to the very low energy levels associated with these periods. These low energy levels produce small objective function values, which means that discrepancies between values quickly become large when considered from a relative perspective.

The variation exhibited by the non-deterministic solutions affirms the necessity of performing multiple simulations for the relevant problem cases, since any one evaluation may converge to a local minimum. Depending on the tolerance placed on optimal solution quality, problem cases that show only small variations in the minima produced over multiple evaluations could be simplified to a single evaluation. For these problem cases, however, it seems difficult to predict when and why pattern search variation 3 and the GA will produce minima that fall within a sufficiently small range for all evaluations.

Objective function 4 clearly produced the most consistent (and therefore possibly the most reliable) results, with all three pattern search variations achieving the same minima for all but one problem case while the best GA minima also shows negligible deviation. The worst minima for pattern search variation 3 also show no deviation apart from that of the aforementioned exception case, while the worst GA minima do show notable deviation in some cases. This indicates that the pattern search approach is very well suited for solving this particular set of optimisation problems, even without the use of a GA search step and multiple evaluations, while the GA is less consistent. Considering these results in comparison to those of the other objective functions, it seems that some objective functions are inherently easier to solve for a global minimum and clearly more suited for a particular optimisation technique.

With regard to simulation runtime, pattern search variation 3 shows significantly shorter average durations than the GA simulations for all problem cases, which can be ascribed to the simplified nature of the GA used in the search step. Both techniques also show an apparently exponential difference in average simulation runtime between site groups 1 and 2 and site group 3, resulting in a much more pronounced difference between the simulation times for pattern search variation 3 and those for the GA in site group 3. This increase in simulation time can be ascribed to the higher number of potential locations analysed in site group 3, since each additional location increases the size of the problem by adding a dimension to the search space. This behaviour indicates that these two methods, especially the GA, may become very expensive when optimising the selected objective functions over a sufficiently large number of locations.

Apart from demonstrating the aforementioned general trends, however, analysis of the average simulation times produced almost no consistent patterns. One exception is the problem cases for objective function 3, which averaged a much longer runtime for the full year cases. This is most likely a result of the more complex approach that objective function 3 has to the full year scenario, where the Megaflex structure is applied to all four seasons separately. With regard to the rest of the results, there is significant variation in the average runtime of different problem cases using the same objective function, the degree of difference between the simulation times of the two methods, as well as the extent to which runtime increases.

After due consideration of all the relevant aspects, this analysis indicates that pattern search variation 3 has the best overall performance for the various types of optimisation problems evaluated in this study and could be used as the sole optimisation method. For all the objective functions apart from objective function 4, however, multiple evaluations are required to guarantee near optimal solution quality, thereby necessitating a corresponding increase in the total simulation runtime. The set of 50 evaluations used throughout this study proved adequate to approximate the best minima for the majority of problem cases, although a higher number of evaluations may further improve the average solution quality at the cost of more time. The total runtime for a set of evaluations ranged from very low (i.e. less than 1 second) to moderate (i.e. approximately 17 minutes) throughout the selected problem cases, although these simulation times would increase proportionally for evaluating a greater number of potential locations.

4.3 Analysis of simulation results

4.3.1 Overview

This section presents and analyses the simulation results of the problem cases analysed for each objective function. As previously stated, only one result is considered for each problem case, namely the distribution that produced the best minimum. Visual representations of the optimal solutions are included, along with the corresponding daily averaged aggregated power profile for each distribution. Since both winter and full year considerations are of particular interest in the context of energy demand, additional aspects of the results are analysed within that seasonal specification.

One aspect considered is the cumulative energy available for a distribution over the pertinent dirunal periods (namely the full day, morning peak and evening peak periods) throughout the full year and the winter months. These cumulative energy values are presented for every problem case that was analysed for either the winter or the full year scenario. For parameter sets where only the summer months and the full year are considered, the cumulative energy values are indicated for the summer scenario as well.

The cumulative available energy at each potential optimisation location is shown for the relevant seasons and diurnal periods in Table 4.6. These values show that Upington receives the highest cumulative annual energy, while Polokwane receives the highest annual energy over the morning peak period and Alexander Bay receives the highest annual energy during the evening peak. When considering only the winter months, Upington and Polokwane still receive the highest cumulative energy throughout the full day and the morning peak period respectively. Throughout the rest of this section, these values are used as a reference for measuring the cumulative energy of optimised distributions relative to the maximum available energy.

Table 4.6: Cumulative available energy [MWh/m^2] for each location.

Location	Cumulative annual energy [MWh/m^2]			Cumulative energy in winter [MWh/m^2]	
	Full day	*Morning peak*	*Evening peak*	*Full day*	*Morning peak*
AB	2552.383	367.715	30.519	544.438	61.377
BFN	2444.946	505.358	3.698	581.101	103.609
DBN	1874.741	434.410	1.219	456.605	97.622
KIM	2335.400	458.863	4.635	538.265	89.801
MID	2416.583	470.093	4.983	561.930	91.463
PKW	2332.189	541.027	1.317	584.422	128.312
PE	2099.824	406.197	5.085	456.543	73.370
PRE	2255.653	504.306	1.964	561.487	115.530
UTN	2558.221	431.345	9.417	592.704	83.588

In addition to the cumulative energy values associated with the results of the specified problem cases, the variability of their corresponding aggregated power profiles over the full seasonal period is also of some interest. The relative standard deviation (RSD) [%] of the cumulative daily energy throughout the specified seasonal period is therefore included as an indication of this variability.

The RSD of the cumulative daily energy at each potential optimisation location is shown for the relevant seasons in Table 4.7. Upington has the lowest RSD for both the full year and winter periods, with Alexander Bay in a relatively close second position while all the other locations show significantly higher variability. Once again, these values are used throughout the rest of this section as a reference for assessing the daily variability of aggregated power profiles for optimised distributions.

Table 4.7: RSD of cumulative daily energy for each location.

Location	**RSD of cumulative daily energy [kWh/m^2]**	
	Full year	***Winter***
AB	18.3	11.4
BFN	24.2	17.4
DBN	32.7	26.4
KIM	26.7	23.5
MID	23.3	18.1
PKW	24.6	17.0
PE	29.4	22.1
PRE	27.5	22.7
UTN	17.0	8.4

For objective functions 4 and 5, which aim to reduce or adjust the variability of cumulative daily energy, auxiliary analyses of the resulting aggregated power and energy profiles are included. The probability distribution of the cumulative daily energy throughout each relevant seasonal period is presented in the form of a histogram normalised relative to the mean value, as well as a histogram normalised relative to the maximum value. Additionally, a Global Tilted Irradiance (GTI) scatter plot showing the hourly distribution of the aggregated power profile throughout the full seasonal period is included.

4.3.2 Objective function 1: Maximisation of the daily averaged energy

4.3.2.1 Full day problem cases

Figures 4.1, 4.3 and 4.5 present the optimal distributions found for site groups 1 to 3 for objective function 1 for the full day, while Figures 4.2, 4.4 and 4.6 present the daily averaged power profiles corresponding to these distributions. For site group 1 the distributions are predominantly allocated to Upington and Alexander Bay, the two locations previously identified as having the highest cumulative energy throughout the year. A seasonal correlation is indicated by the fact that 100 % capacity is allocated to Alexander Bay for the spring and summer cases, while most of the capacity for the winter and autumn cases is allocated to Upington, which has a slightly more northerly latitude.

The daily averaged GTI profiles indicate a similar peak power level for all seasons, although the variation during the early morning and late afternoon hours is more significant. The autumn, winter and full year profiles are highly similar, which can be ascribed to their similar distributions, and are spread relatively symmetrically around the midday point. There is a clear discrepancy, however, between the spring and summer profiles, with summer showing significant lower power levels in the early morning hours. Interestingly, the spring profile has the lowest peak value of all the profiles, yet shows the highest power levels during both the late afternoon and especially the early morning hours.

The distributions for site group 2 show a more clear seasonal trend, with the generation capacity predominantly allocated to the second-most southern location for the spring and summer cases, the central location in terms of latitude for the autumn and full year cases, and the northernmost location for the winter case. It is not surprising that no capacity is allocated to Port Elizabeth or Durban for any of the problem cases, since these two locations offer the lowest cumulative energy within the site group.

When considering the GTI profiles, the difference between seasons for the morning and late afternoon hours is less pronounced than for site group 1, though the spring profile clearly favours the morning while the summer profile is the highest during the late afternoon. In contrast to site group 1, the spring profile also shows the highest peak value of all five profiles. This spring peak is slightly higher than its counterpart for site group 1, though resulting in notably lower power levels during the late afternoon, while the other profiles all show lower peak values than the equivalent profiles for site group 1.

The generally lower power profiles for site group 2 can be ascribed to the fact that the two dominant locations in site group 1 offer higher overall energy levels than any of the locations in site group 2. It is therefore not unexpected that the distributions for site group 3, which encompasses all the potential locations for site groups 1 and 2, largely correspond to the results for site group 1. There are small discrepancies between the group 1 and 3 distributions for the autumn and full year cases, though the GTI profiles are not notably different. This either indicates that slight variations in the distribution could produce equivalent optimisation minima for this objective function, or that the complexity introduced by a larger selection of potential locations can impact the solution quality.

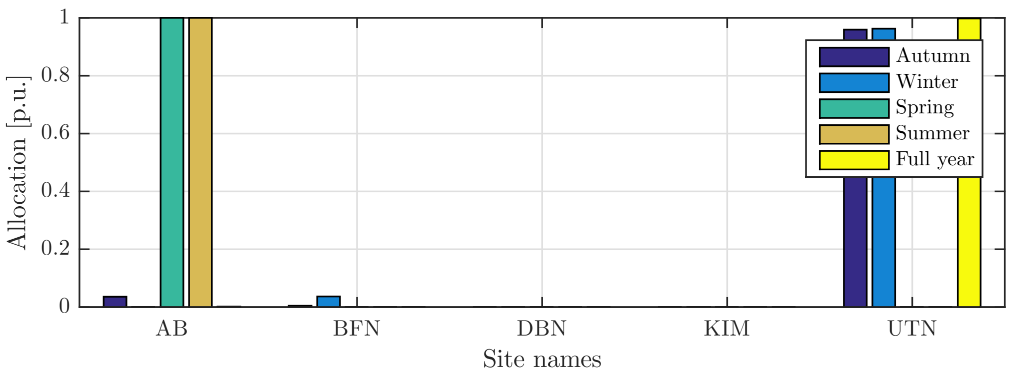

Figure 4.1: Optimal distributions in site group 1 for objective function 1 for the full day during each season.

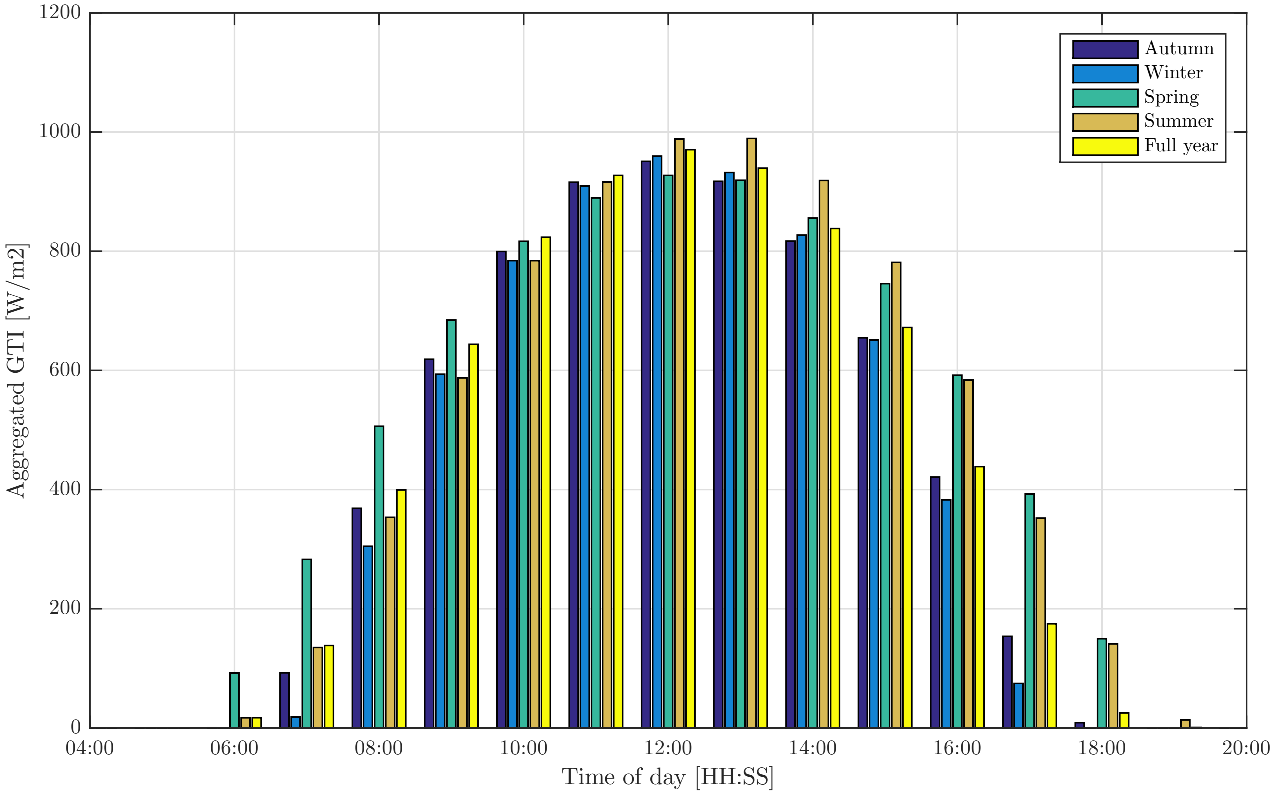

Figure 4.2: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 1 for the full day during each season.

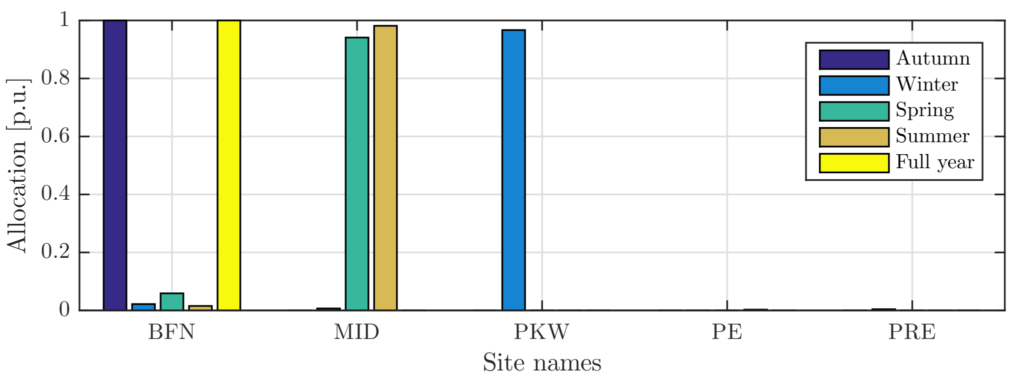

Figure 4.3: Optimal distributions in site group 2 for objective function 1 for the full day during each season.

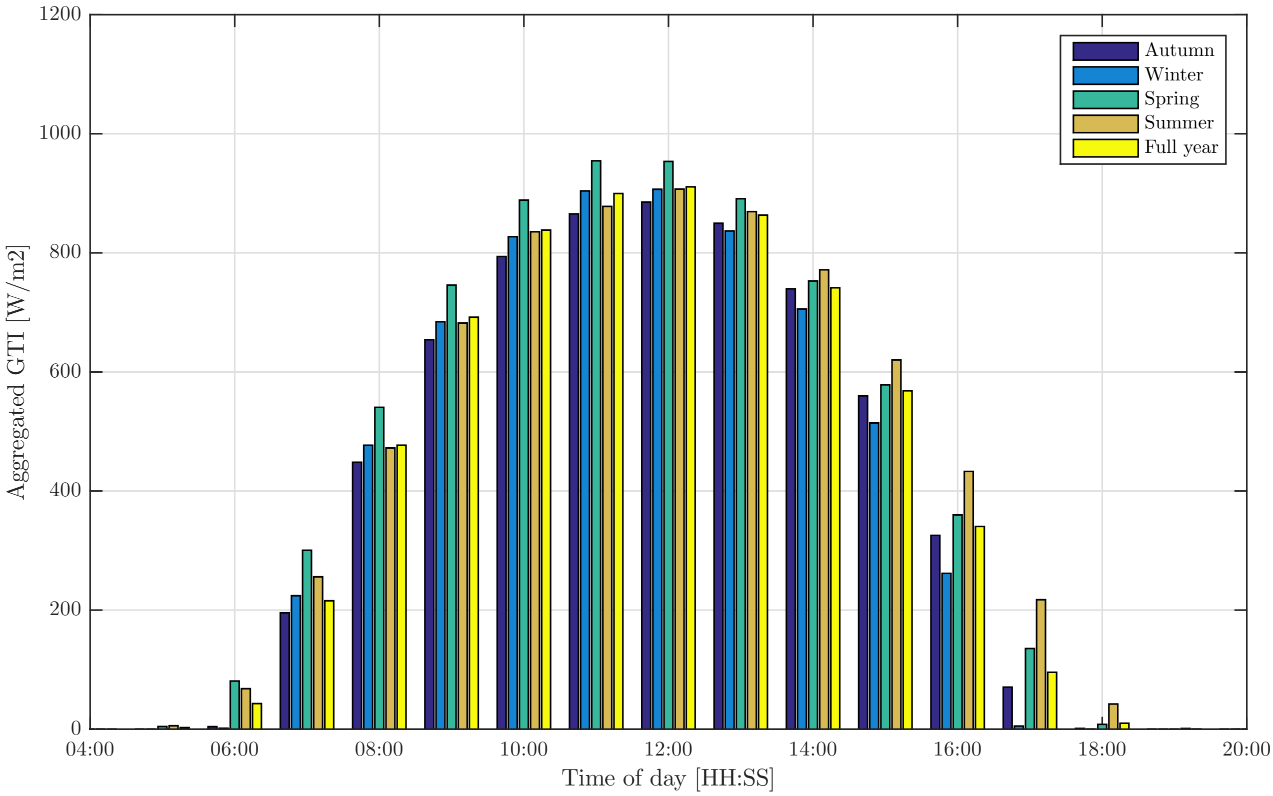

Figure 4.4: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 1 for the full day during each season.

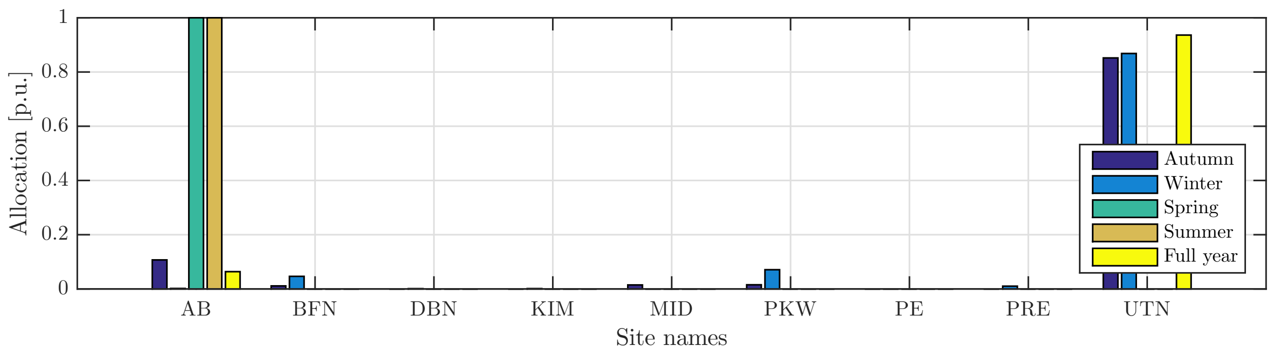

Figure 4.5: Optimal distributions in site group 3 for objective function 1 for the full day during each season.

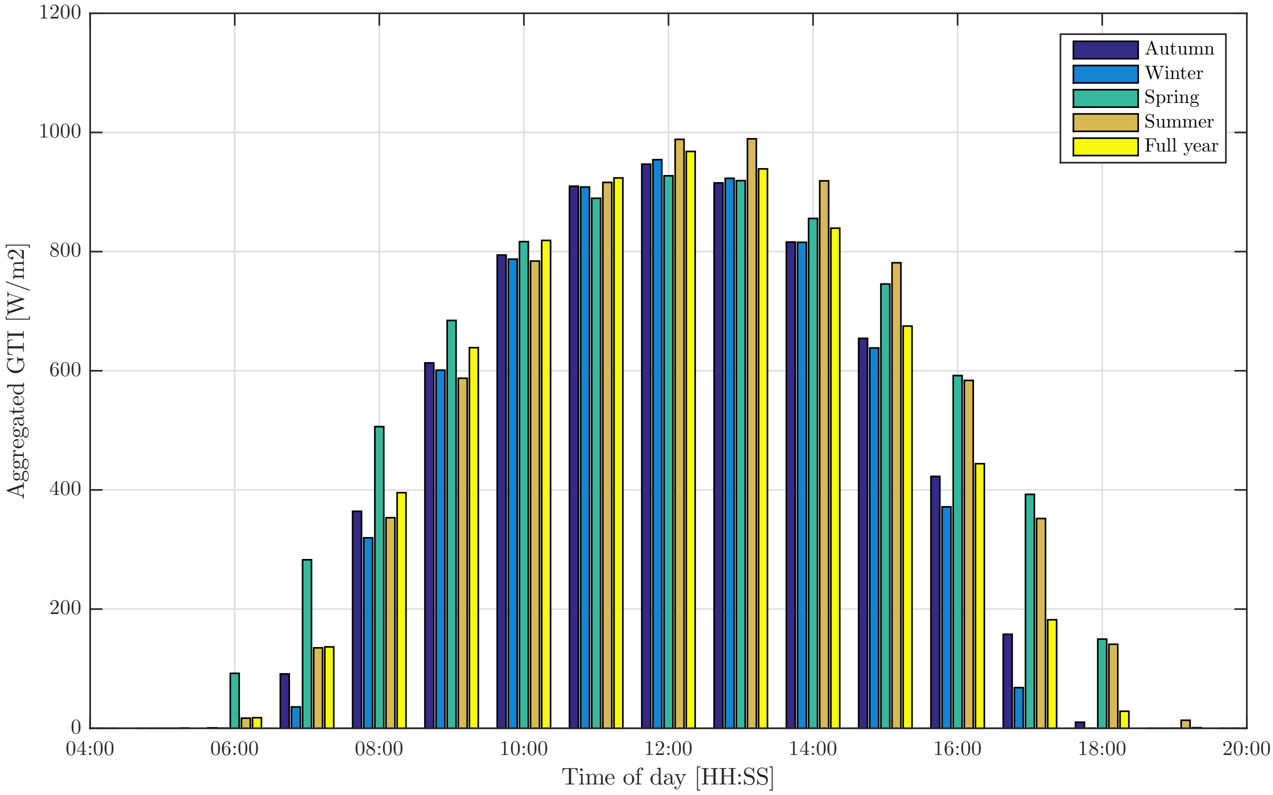

Figure 4.6: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 1 for the full day during each season.

4.3.2.2 Morning peak problem cases

Figures 4.7, 4.9 and 4.11 present the optimal distributions found for site groups 1 to 3 for objective function 1 for the morning peak period, while Figures 4.8, 4.10 and 4.12 present the daily averaged power profiles corresponding to these distributions. The results for site group 1 present no seasonal variation, with 100 % capacity allocated to the second-most eastern location for all problem cases. Although Durban is located at a significantly more eastern longitude, it is not unexpected that this location is avoided since it has significantly lower cumulative energy levels, both during the full day and the morning peak period, than any other potential location.

When comparing the daily averaged GTI profiles for these distributions to those produced by the results for the full day scenario, the power levels during the morning peak period are notably higher for all seasons. In accordance with this shift, the power levels during the late afternoon hours are notably lower than for the full day results, while the peak power values are also diminished for all but the spring profile.

The results for site group 2 favour the two locations with the highest cumulative energy levels for the morning peak period, with the total capacity allocated to the northernmost location for the autumn, winter and full year cases, and to the middle location in terms of latitude for the summer and spring cases. The difference between the daily averaged GTI profiles for the morning peak period and those corresponding to the full day distributions is not as pronounced as for site group 1, though clear improvements can be observed for all but the winter case. Once again, these adjustments also result in lower peak values and power levels during the late afternoon.

The distributions found for site group 3 largely correspond to the results for site group 2, although there are once again small discrepancies that do not present significant changes in the daily averaged GTI profiles. The difference in power levels when compared to the site group 3 profiles for the full day scenario, however, is quite evident for every season.

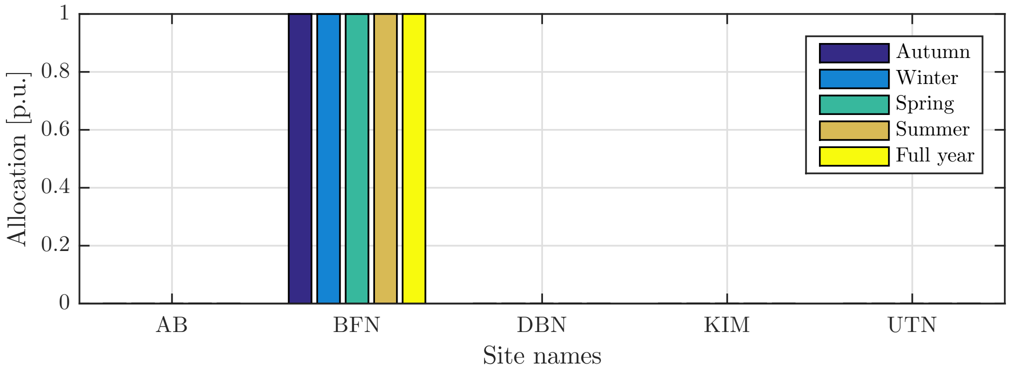

Figure 4.7: Optimal distributions in site group 1 for objective function 1 for the morning peak period during each season.

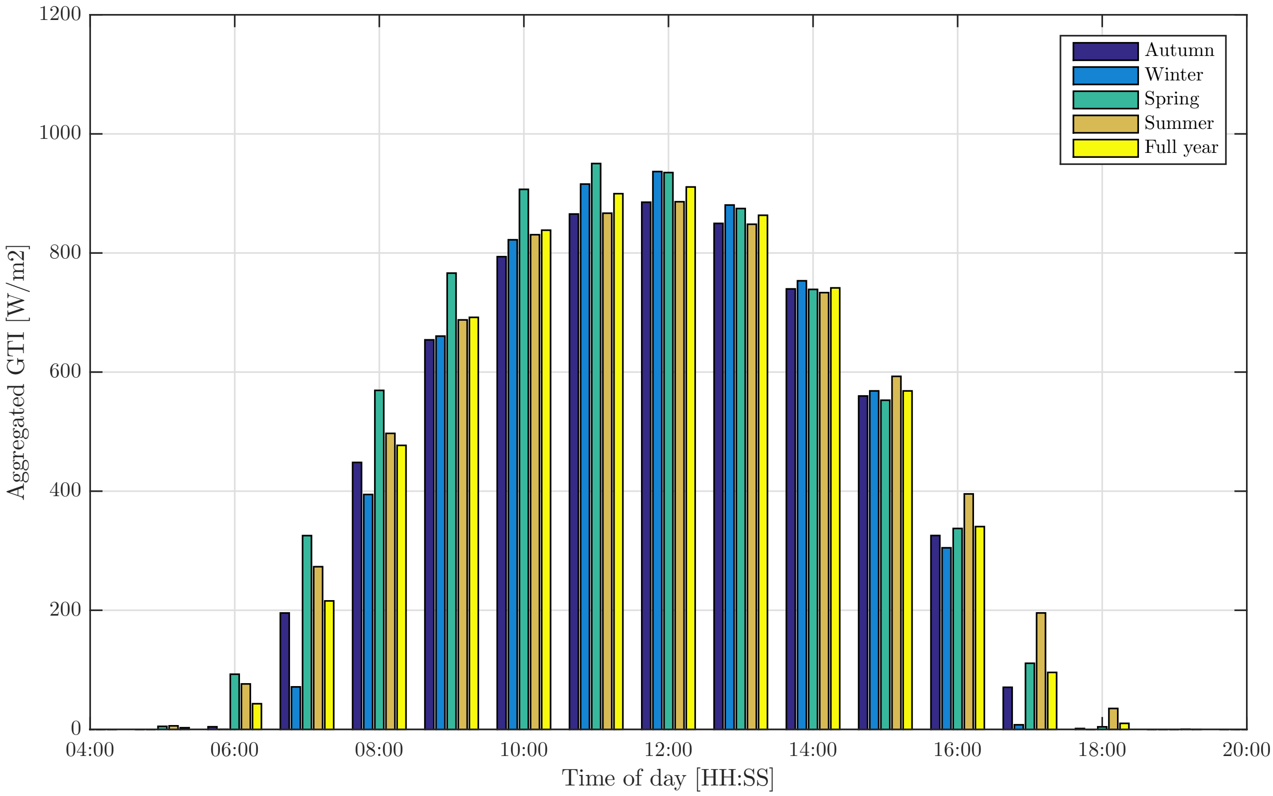

Figure 4.8: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 1 for the morning peak period during each season.

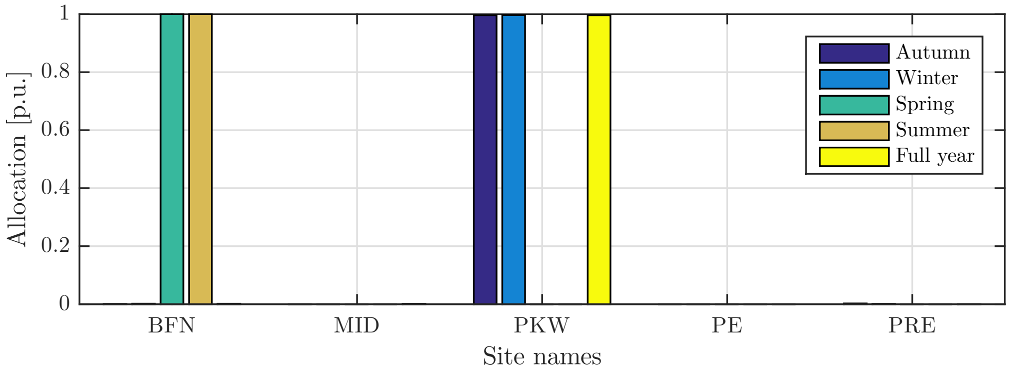

Figure 4.9: Optimal distributions in site group 2 for objective function 1 for the morning peak period during each season.

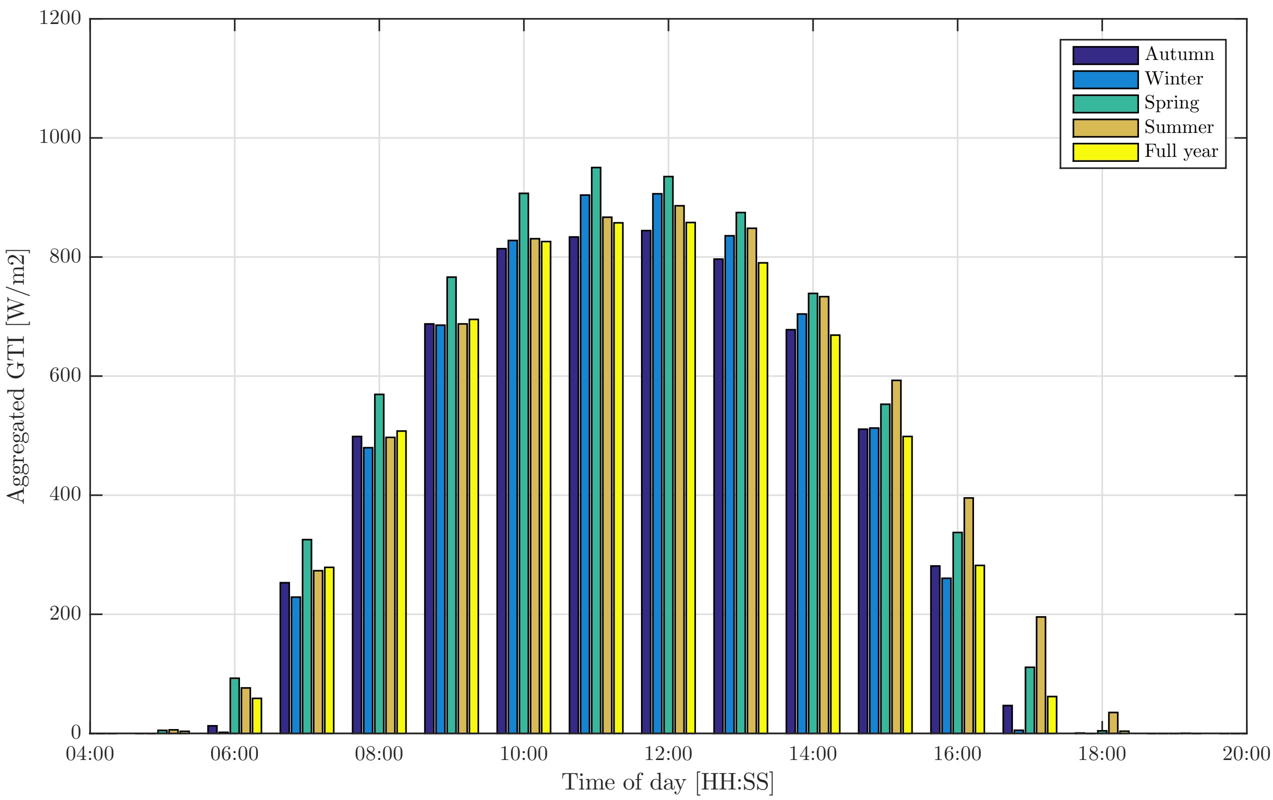

Figure 4.10: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 1 for the morning peak period during each season.

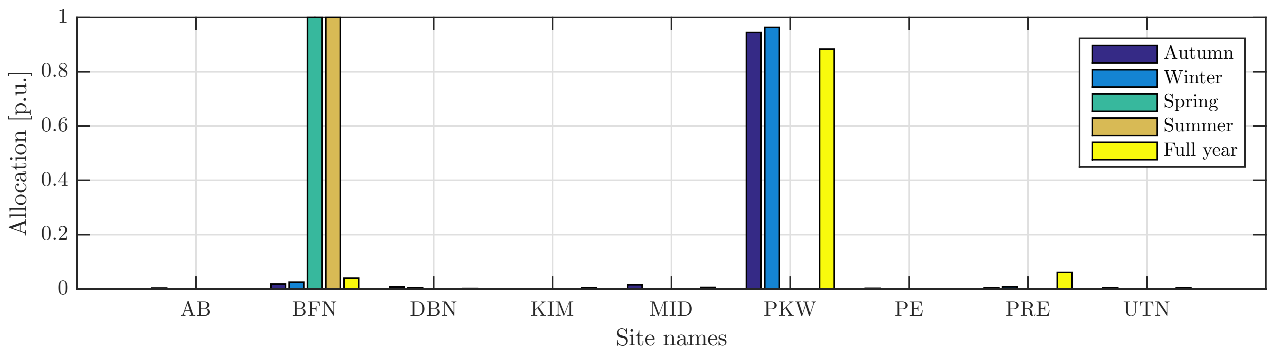

Figure 4.11: Optimal distributions in site group 3 for objective function 1 for the morning peak period during each season.

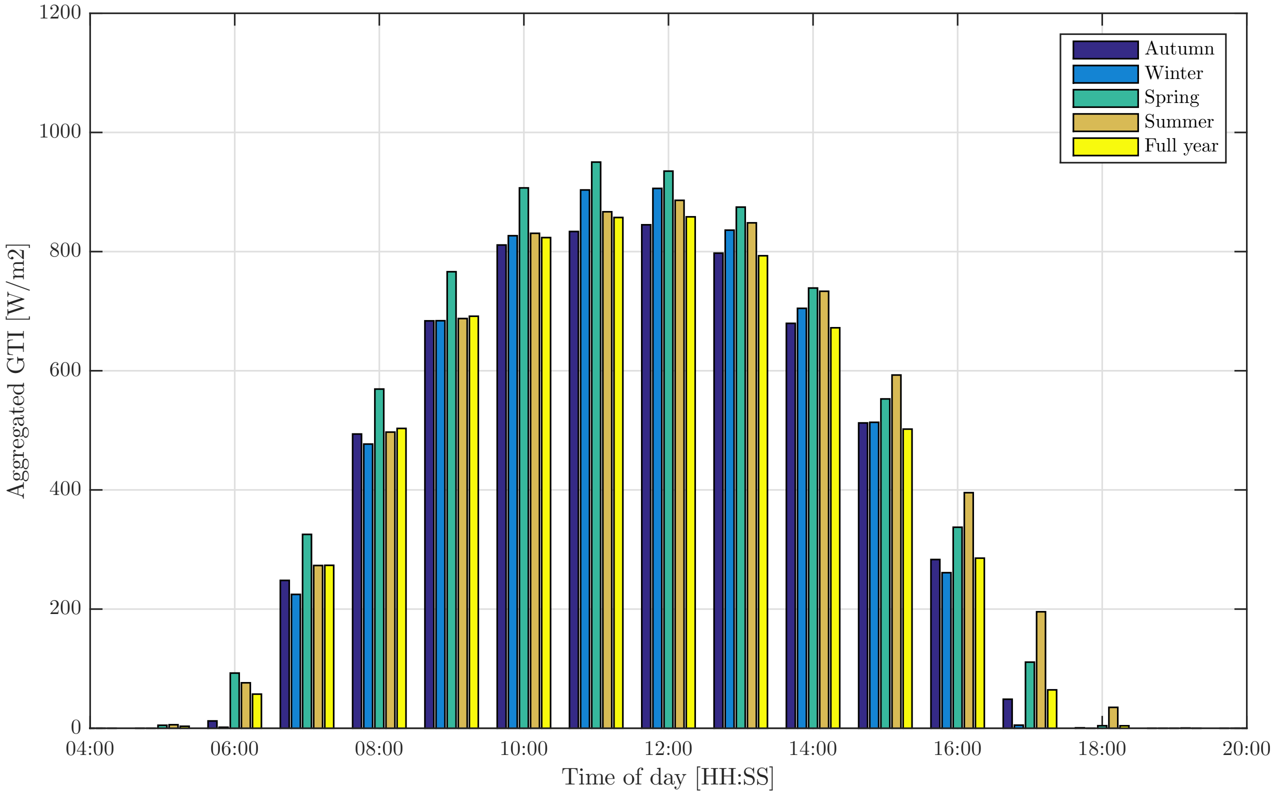

Figure 4.12: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 1 for the morning peak period during each season.

4.3.2.3 Evening peak problem cases

Figures 4.13, 4.15 and 4.17 present the optimal distributions found for site groups 1 to 3 for objective function 1 for the evening peak period, while Figures 4.14, 4.16 and 4.18 present the daily averaged power profiles corresponding to these distributions. As expected, the full capacity is allocated to the western-most location for the summer and full year cases for site group 1. The daily averaged GTI profiles show distinctly higher power levels during the late afternoon compared to the equivalent profiles produced for the morning peak period as well as the full year profile produced for the full day scenario (the distributions for the evening peak and full day summer cases are the same). Even though there is no significant level of energy available during the evening peak period itself, using it as an optimisation parameter succeeded in prioritising the availability of energy during the late afternoon.

Interestingly, the results for site group 2 favour the southernmost location, which results in GTI profiles with slightly lower power levels during the late afternoon than the equivalent profiles produced for either the full day or morning peak cases. This indicates that sufficiently high and/or differing evening peak energy levels are required within the available locations to ensure results with increased power levels during the late afternoon. In view of the unfavourable solutions found for site group 2, it is unsurprising that the results for site group 3 are identical to those produced for site group 1.

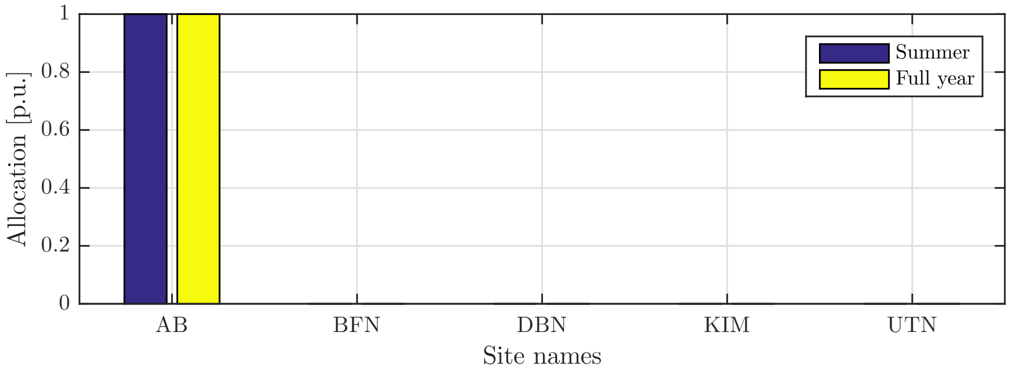

Figure 4.13: Optimal distributions in site group 1 for objective function 1 for the evening peak period during each season.

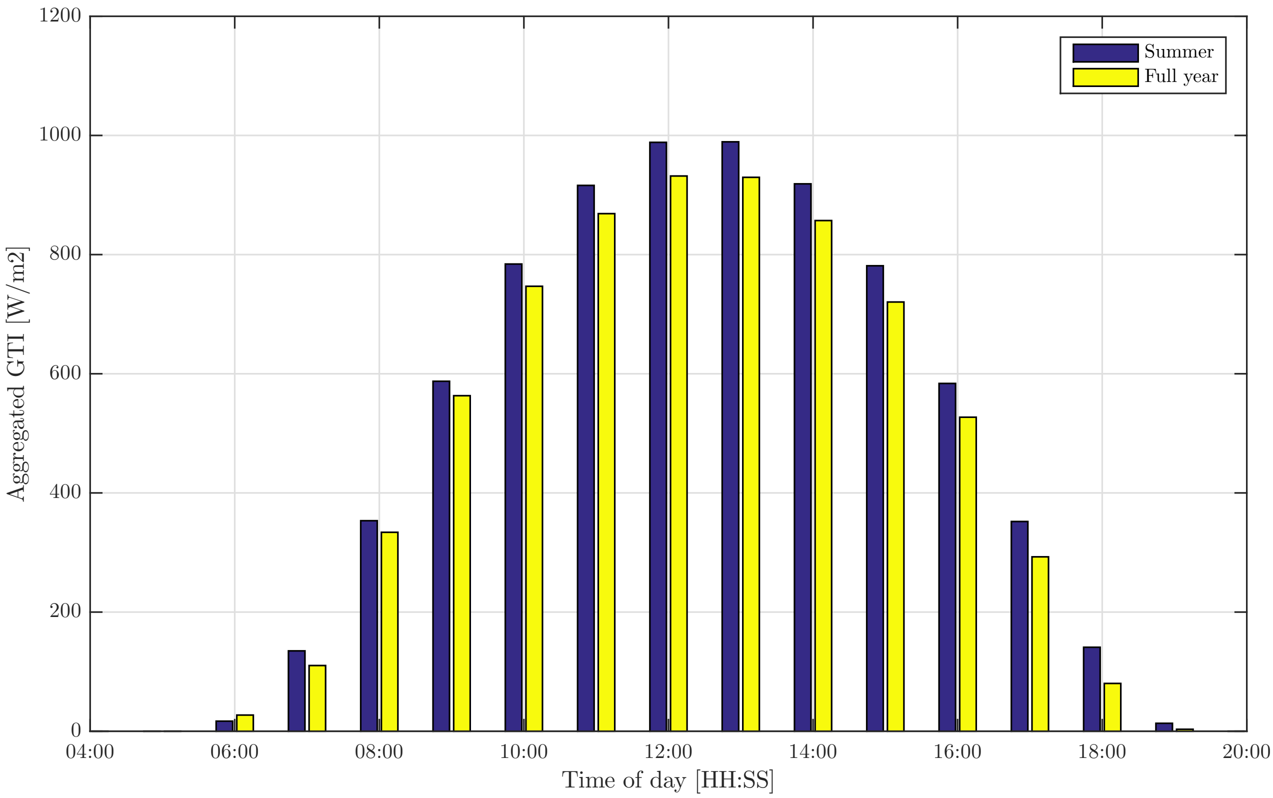

Figure 4.14: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 1 for the evening peak period during each season.

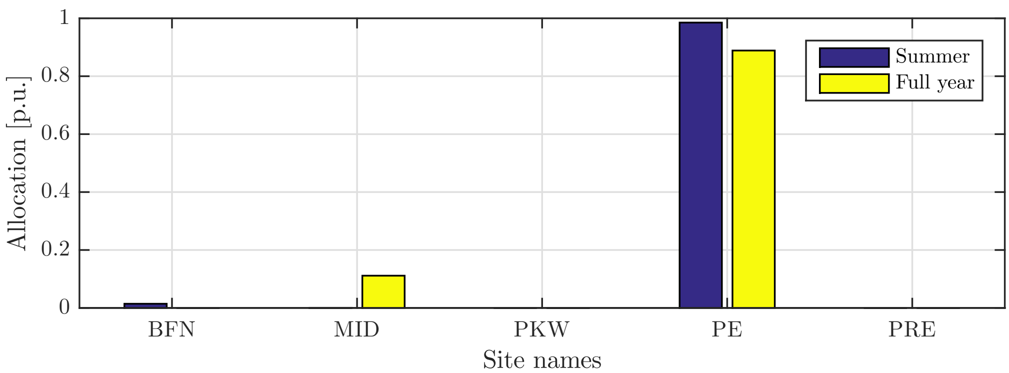

Figure 4.15: Optimal distributions in site group 2 for objective function 1 for the evening peak period during each season.

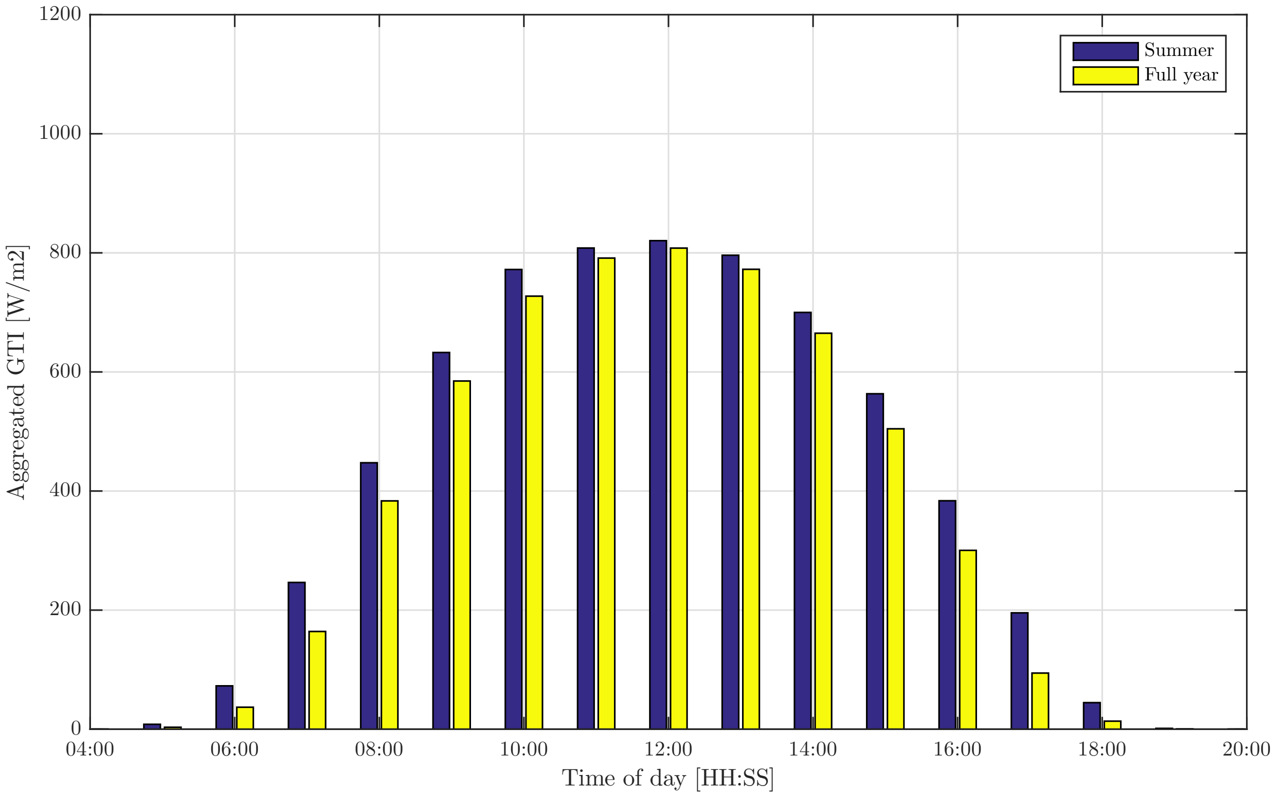

Figure 4.16: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 1 for the evening peak period during each season.

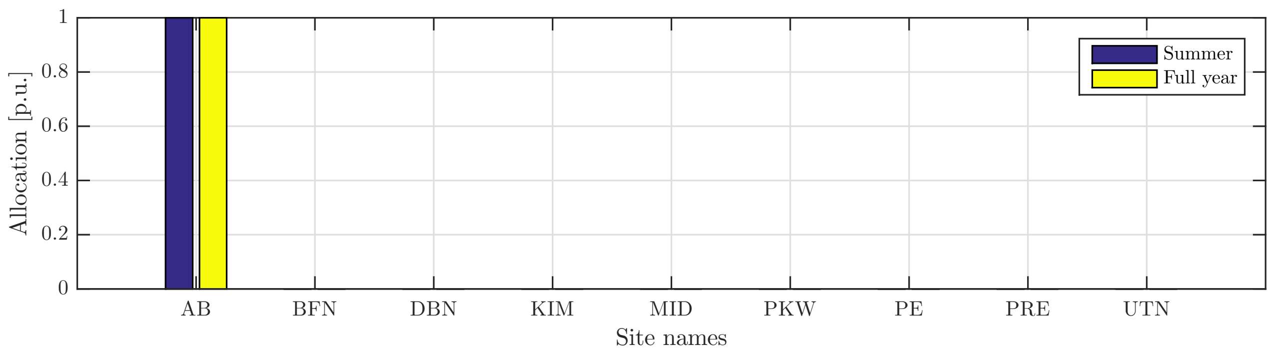

Figure 4.17: Optimal distributions in site group 3 for objective function 1 for the evening peak period during each season.

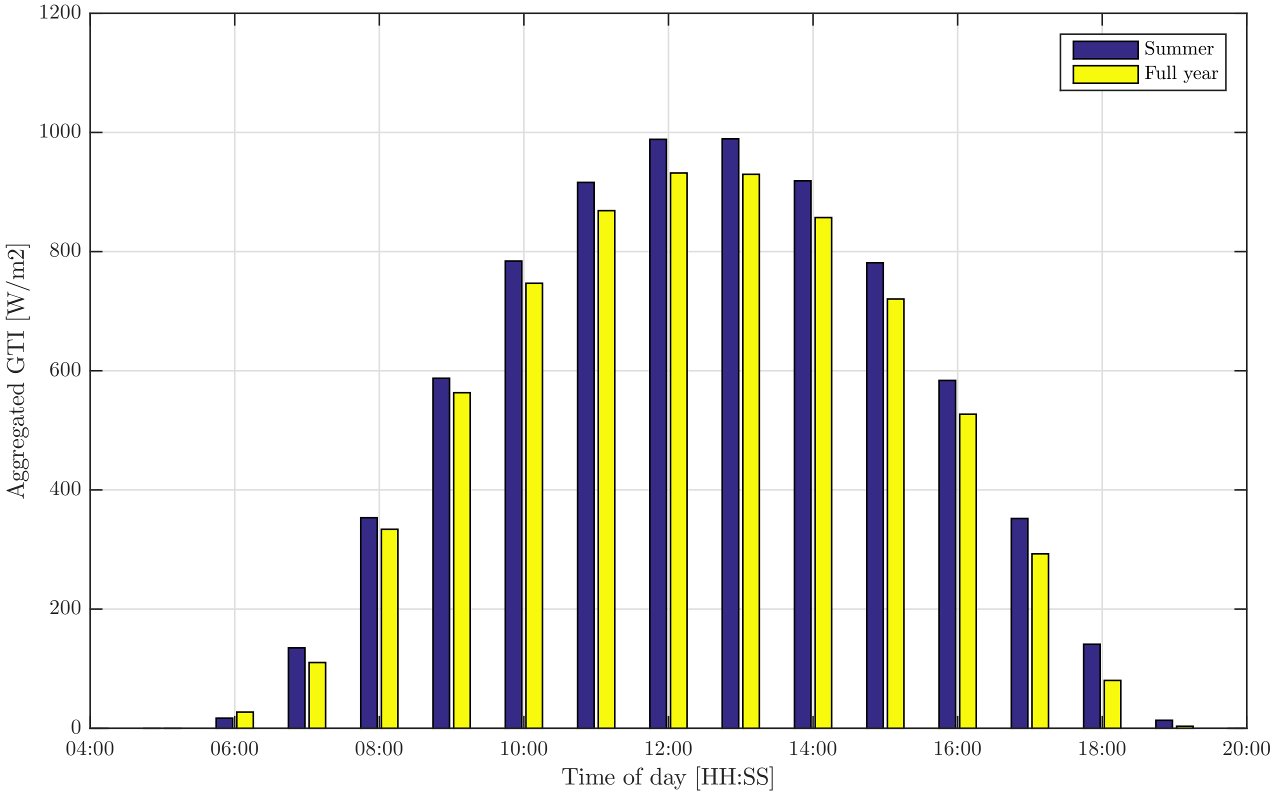

Figure 4.18: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 1 for the evening peak period during each season.

4.3.2.4 Summary

In terms of distribution, the optimal solutions for problem cases analysed with objective function 1 consistently allocate all or most of the available capacity to a single location. This makes sense in the context of the objective function, since the maximisation of daily averaged energy for various diurnal periods is clearly correlated to the single locations exhibiting the highest cumulative energy levels for each set of simulation parameters.

The results of the problem cases analysed for site group 1 confirm the correlation of diurnal and longitudinal variation with regard to maximising solar energy, with more eastern locations providing higher power levels in the early morning hours and more western locations increasing power delivery during the late afternoon. The results for site group 2, on the other hand, confirm the correlation of seasonal and latitudinal variation, with more northern locations favoured for winter and autumn, and more southern locations favoured for spring and summer.

Comparative analysis of the three site groups show that the results for site group 3 generally replicate or closely resemble the best results found in either site group 1 or 2 for each problem case. As previously stated, the discrepancies occurring for some problem cases could either indicate that the true optimal solution becomes more difficult to find due to an increase in search parameters or that the additional search parameters create an area in the search space where small variations can occur with negligible change in the objective function fitness value.

Table 4.8 presents the cumulative energy values associated with the optimal distributions found for relevant problem cases analysed with objective 1, while Table 4.9 presents the variability of the cumulative daily energy associated with each solution. For the distributions consisting of a single location allocated with 100 % capacity, the cumulative energy and RSD variables are equivalent to the values presented in Tables 4.6 and 4.7.

Comparative analysis of the cumulative energy values show that the diurnal period as well as the season used for optimisation is significant: most of the distributions optimised for energy output during the morning peak period deliver significantly less energy annually than the equivalent distributions for the full year, while some distributions optimised for the winter (e.g. the full day case for site group 2) have distinctly lower annual energy levels than their full year counterparts. In contrast, it can be observed that even the low degree of distribution exhibited for these solutions can improve its yearly or seasonal variability from that of the least variable single location. Examples include the full day winter cases for site groups 1, 2 and 3, the full day full year case for site group 3, and the morning peak winter and full year cases for site group 2.

Table 4.8: Cumulative available energy for problem cases analysed for objective function 1 with GTI profiles.

Problem case description			**Cumulative annual energy [MWh/m^2]**			**Cumulative energy in winter [MWh/m^2]**	
Diurnal period	***Site group***	***Seasonal period***	***Full day***	***Morning peak***	***Evening peak***	***Full day***	***Morning peak***
Full day	1	Winter	2553.930	434.090	9.204	592.249	84.330
		Full year	2558.210	431.225	9.457	592.614	83.546
	2	Winter	2334.923	539.595	1.397	584.095	127.460
		Full year	2444.946	505.358	3.698	581.101	103.609
	3	Winter	2532.802	443.224	8.524	590.953	87.997
		Full year	2557.849	427.290	10.762	589.629	82.172
Morning peak	1	Winter	2444.775	505.337	3.697	581.064	103.607
		Full year	2444.946	505.358	3.698	581.101	103.609
	2	Winter	2332.311	540.916	1.322	584.388	128.250
		Full year	2332.511	540.834	1.327	584.372	128.203
	3	Winter	2332.570	539.426	1.381	583.649	127.475
		Full year	2332.112	535.908	1.516	582.212	125.914
Evening peak	1	Summer	2552.383	367.715	30.519	544.438	61.377
		Full year	2552.383	367.715	30.519	544.438	61.377
	2	Summer	2104.920	407.651	5.065	458.379	73.812
		Full year	2135.089	413.311	5.073	468.276	75.384
	3	Summer	2552.383	367.715	30.519	544.438	61.377
		Full year	2552.383	367.715	30.519	544.438	61.377

Table 4.9: RSD of cumulative daily energy for problem cases analysed for objective function 1 with GTI profiles.

Problem case description			RSD of cumulative daily energy [kWh/m^2]	
Diurnal period	***Site group***	***Seasonal period***	***Full year***	***Winter***
Full day	1	Winter	16.5	8.2
		Full year	17.0	8.4
	2	Winter	23.8	16.5
		Full year	24.2	17.4
	3	Winter	15.1	7.9
		Full year	16.3	8.3
Morning peak	1	Winter	24.2	17.4
		Full year	24.2	17.4
	2	Winter	24.5	16.9
		Full year	24.5	16.9
	3	Winter	23.7	16.4
		Full year	21.9	15.4
Evening peak	1	Summer	18.3	11.4
		Full year	18.3	11.4
	2	Summer	28.9	21.7
		Full year	26.1	19.9
	3	Summer	18.3	11.4
		Full year	18.3	11.4

4.3.3 Objective function 2: Minimisation of the CV for the daily averaged power profile

4.3.3.1 Full day problem cases

Figures 4.19, 4.21 and 4.23 present the optimal distributions found for site groups 1 to 3 for objective function 2 for the full day, while Figures 4.20, 4.22 and 4.24 present the daily averaged power profiles corresponding to these distributions. In contrast to the results for objective function 1, the distributions for site group 1 show a significant division of the total capacity for all except the spring case. For the winter, summer and full year cases the distribution is split between the westernmost and easternmost locations, while for the autumn case it is divided between the westernmost location and the location that is central with regard to longitude.

Apart from the spring case, the GTI profiles for these distributions show much lower peak power values than the equivalent profiles for the full day problem cases optimised with objective function 1. In the context of this lower peak power level and compared to the aforementioned equivalent profiles, certain profiles appear somewhat adjusted towards the diurnal extremities. The summer profile is shifted towards the morning peak period, while the full year profile is shifted towards the evening peak period. The profile for the spring distribution, which is the same as for the full day problem case optimised with objective function 1, remains perhaps the most balanced profile with regard to both the morning and evening peak considerations.

In contrast to site group 1, the results for site group 2 allocate most or all of the capacity to a single location for every problem case. This can probably be ascribed to the lack of longitudinal variation within site group 2 that would enable a more diurnally balanced distribution. With regard to seasonal variation, a clear trend can be observed in accordance with the latitudinal correlation established in the previous section. Due to very similar distributions, the winter and spring GTI profiles closely resemble the equivalent profiles for the full day problem cases optimised with objective function 1. For the full year, autumn and summer cases, however, the GTI profiles exhibit distinctly lower peak power values, with the full year and autumn cases also showing a shift towards the morning peak period.

The results for site group 3 retain the summer and spring distributions of site group 1, with a different set of longitudinally dispersed distributions for winter, autumn and full year cases. Once again, the GTI profiles for all but the spring distribution show lower peak power values than the equivalent profiles for the full day cases optimised with objective function 1. The summer, autumn and winter profiles also show a shift towards the morning peak.

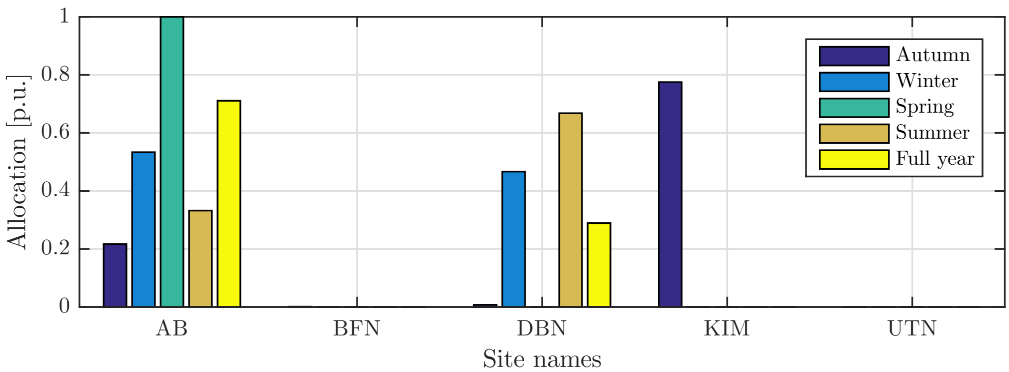

Figure 4.19: Optimal distributions in site group 1 for objective function 2 for the full day during each season.

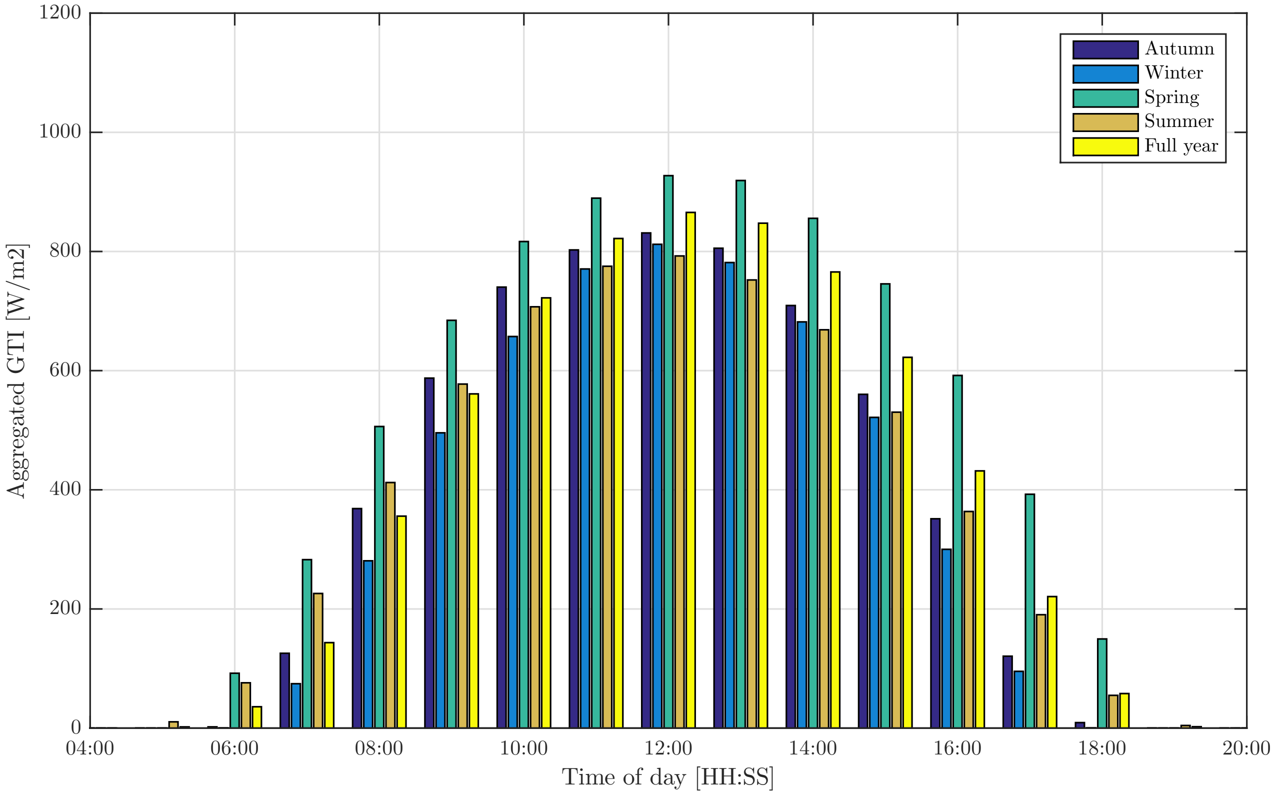

Figure 4.20: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 2 for the full day during each season.

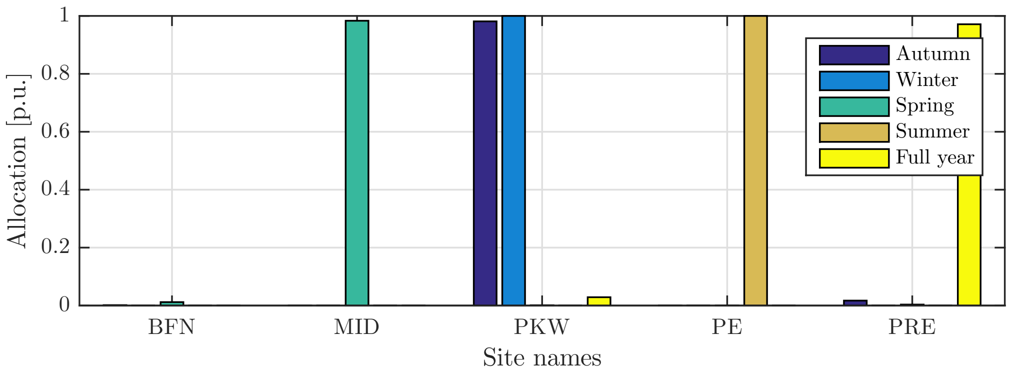

Figure 4.21: Optimal distributions in site group 2 for objective function 2 for the full day during each season.

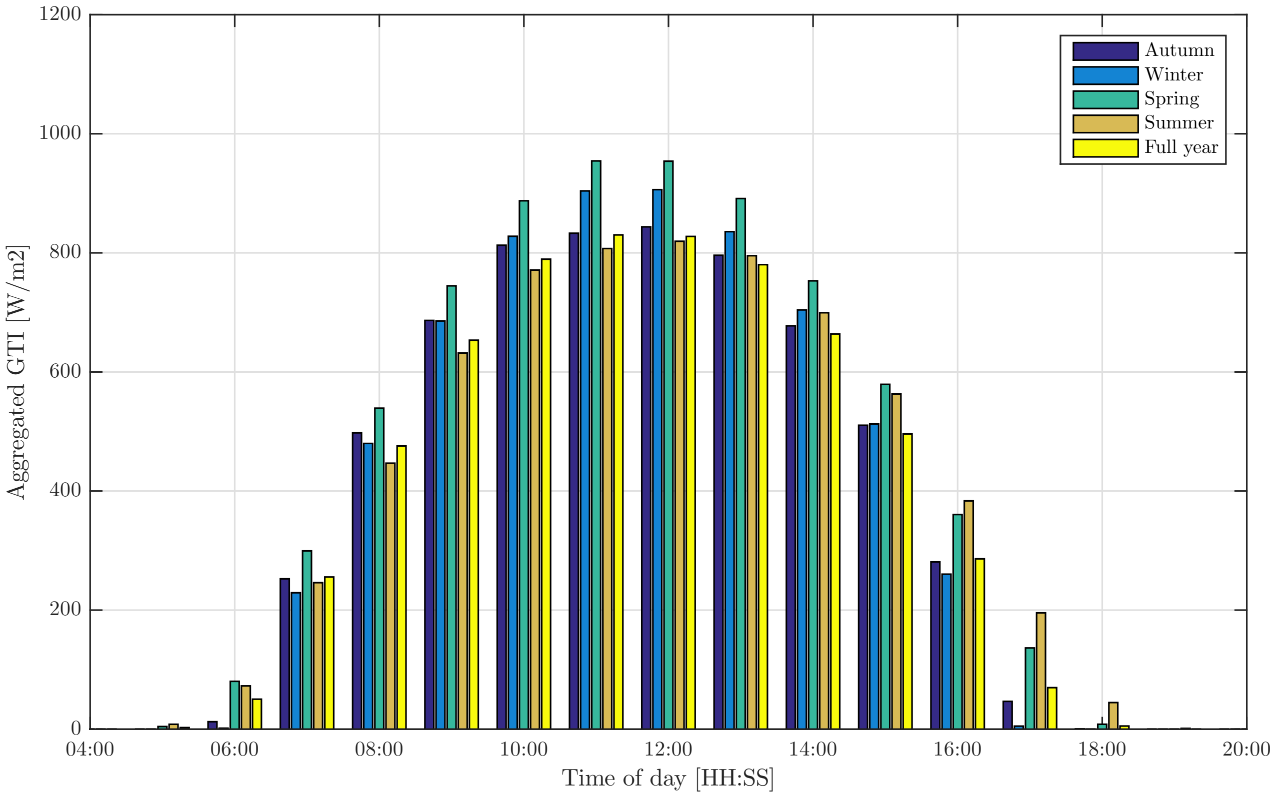

Figure 4.22: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 2 for the full day during each season.

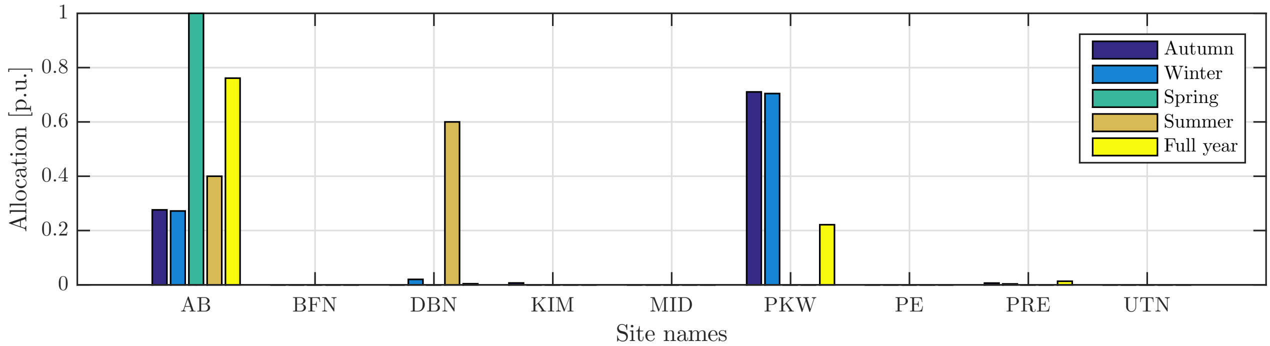

Figure 4.23: Optimal distributions in site group 3 for objective function 2 for the full day during each season.

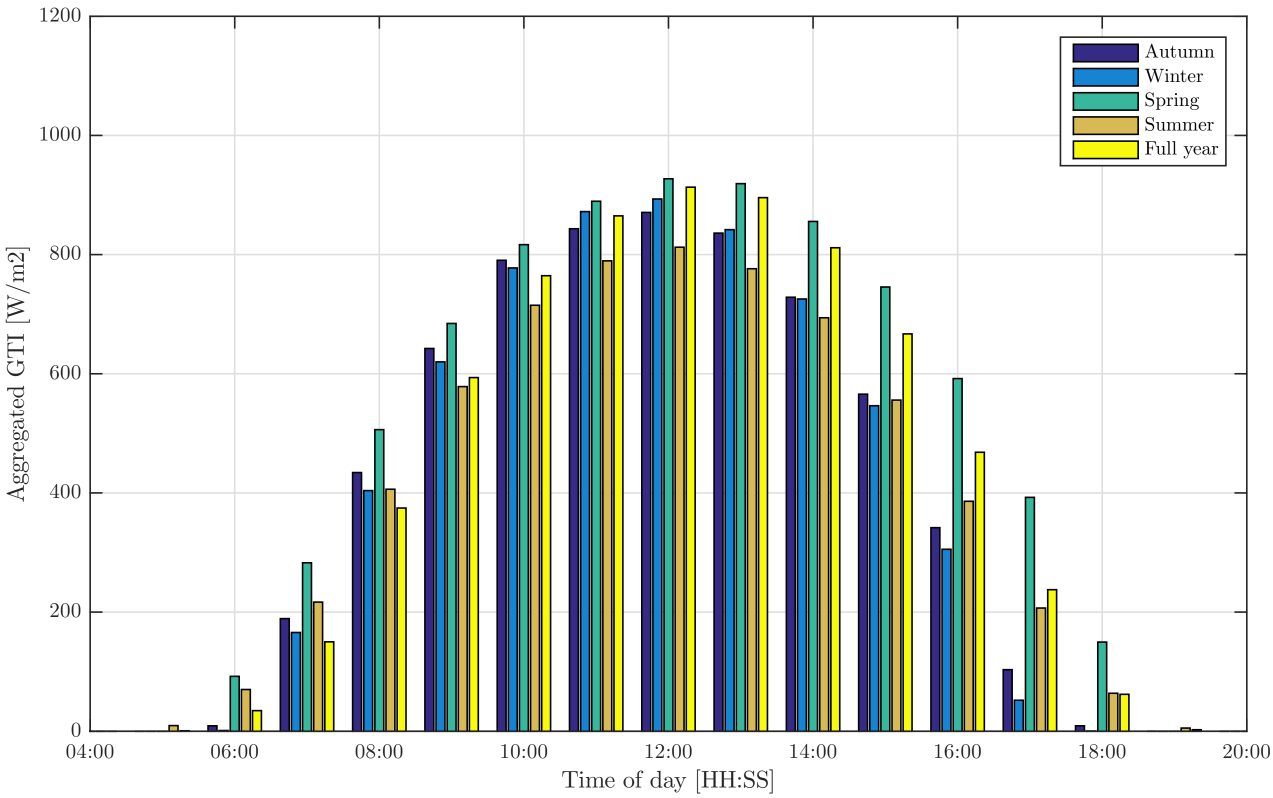

Figure 4.24: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 2 for the full day during each season.

4.3.3.2 Combined peak problem cases

Figures 4.25, 4.27 and 4.29 present the optimal distributions found for site groups 1 to 3 for objective function 2 for the full day, while Figures 4.26, 4.28 and 4.30 present the daily averaged power profiles corresponding to these distributions. Similar to the results for several of the full day problem cases, the distributions for site group 1 are divided between the easternmost and westernmost locations, albeit with different ratios. The resulting GTI profile for the full year case has a lower peak power value and is shifted towards the morning peak with regard to the equivalent profile for the full day problem case. The summer profile, on the other hand, has a higher peak power value and is shifted towards the evening peak with regard to the equivalent profile for the full day scenario.

For site group 2, the distribution for the summer case matches the result for the equivalent full day case. The full year distribution, however, is allocated to the northernmost location, resulting in slightly higher peak power levels and slightly lower power levels during the late afternoon compared to the equivalent profile for the full day case. The distributions for site group 1 are replicated for site group 3, resulting in very similar differences of peak power values and diurnal shift with regard to the equivalent full day power profiles.

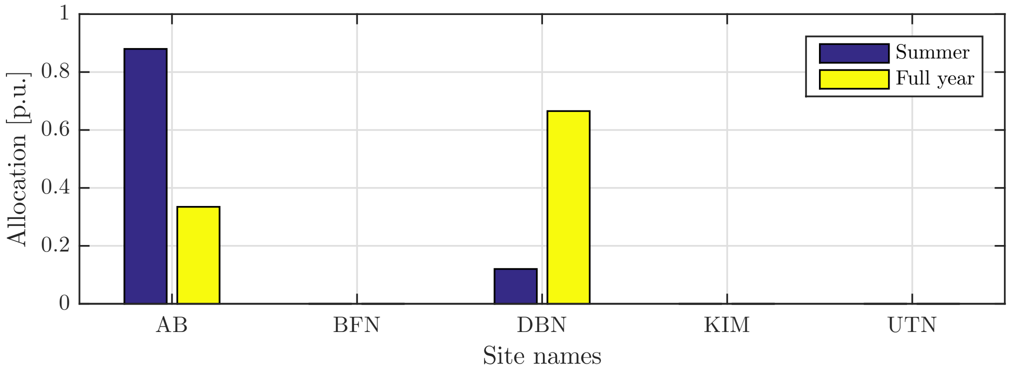

Figure 4.25: Optimal distributions in site group 1 for objective function 2 for the combined peak periods during each season.

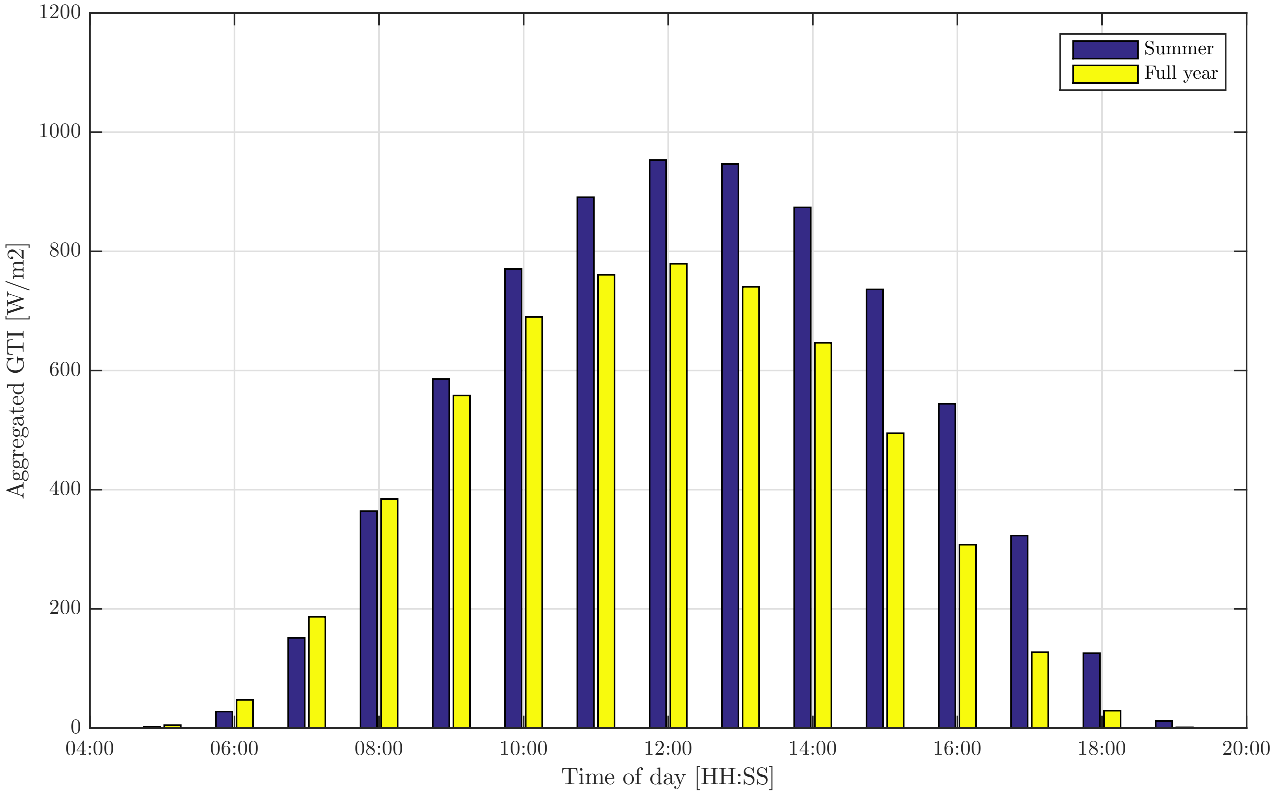

Figure 4.26: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 2 for the combined peak periods during each season.

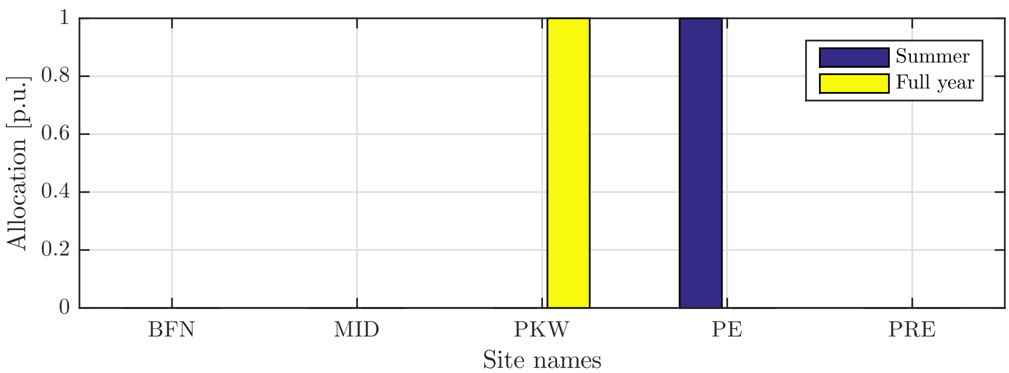

Figure 4.27: Optimal distributions in site group 2 for objective function 2 for the combined peak periods during each season.

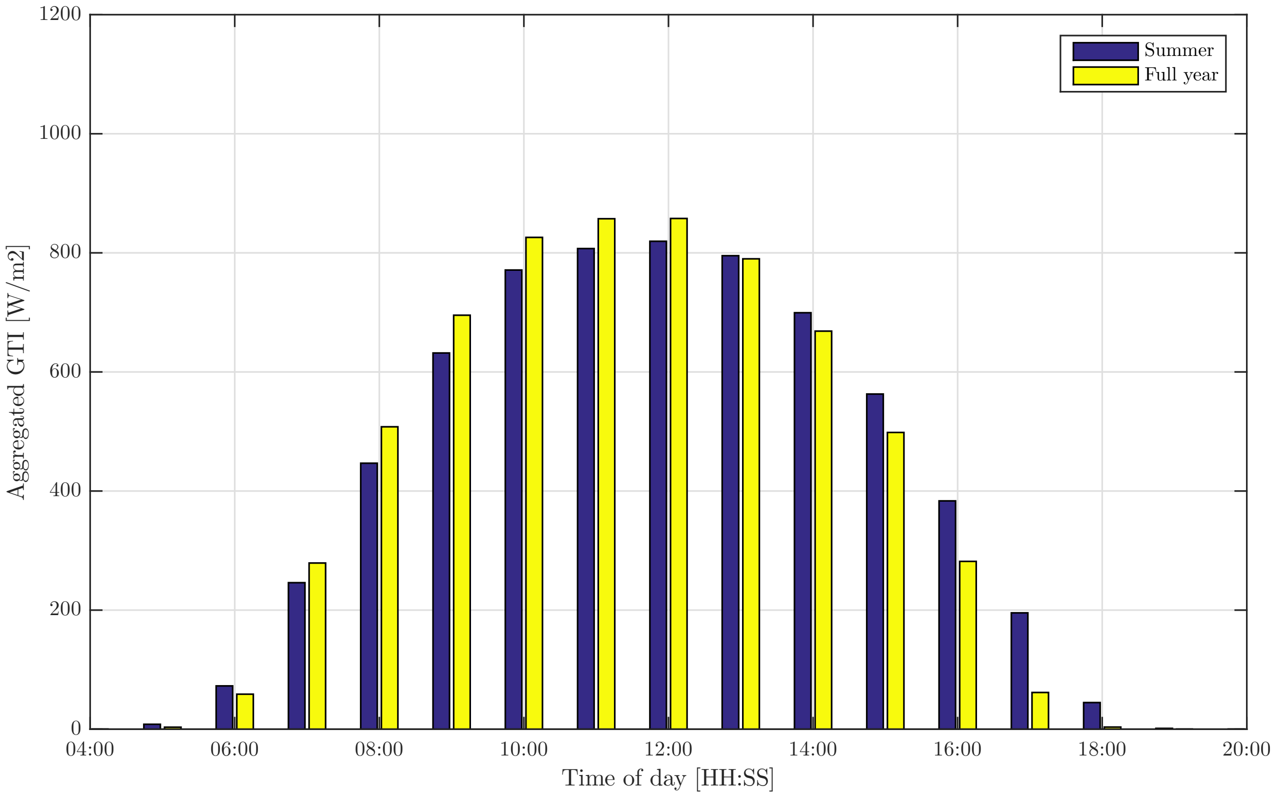

Figure 4.28: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 2 for the combined peak periods during each season.

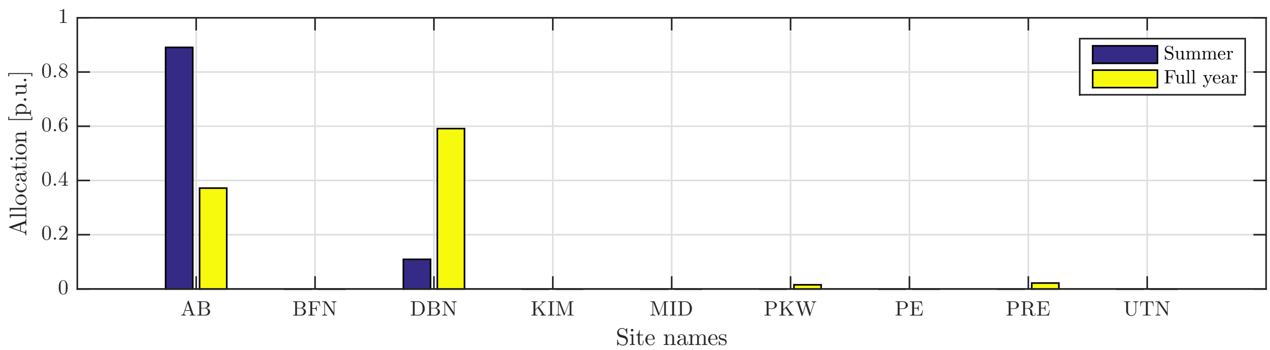

Figure 4.29: Optimal distributions in site group 3 for objective function 2 for the combined peak periods during each season.

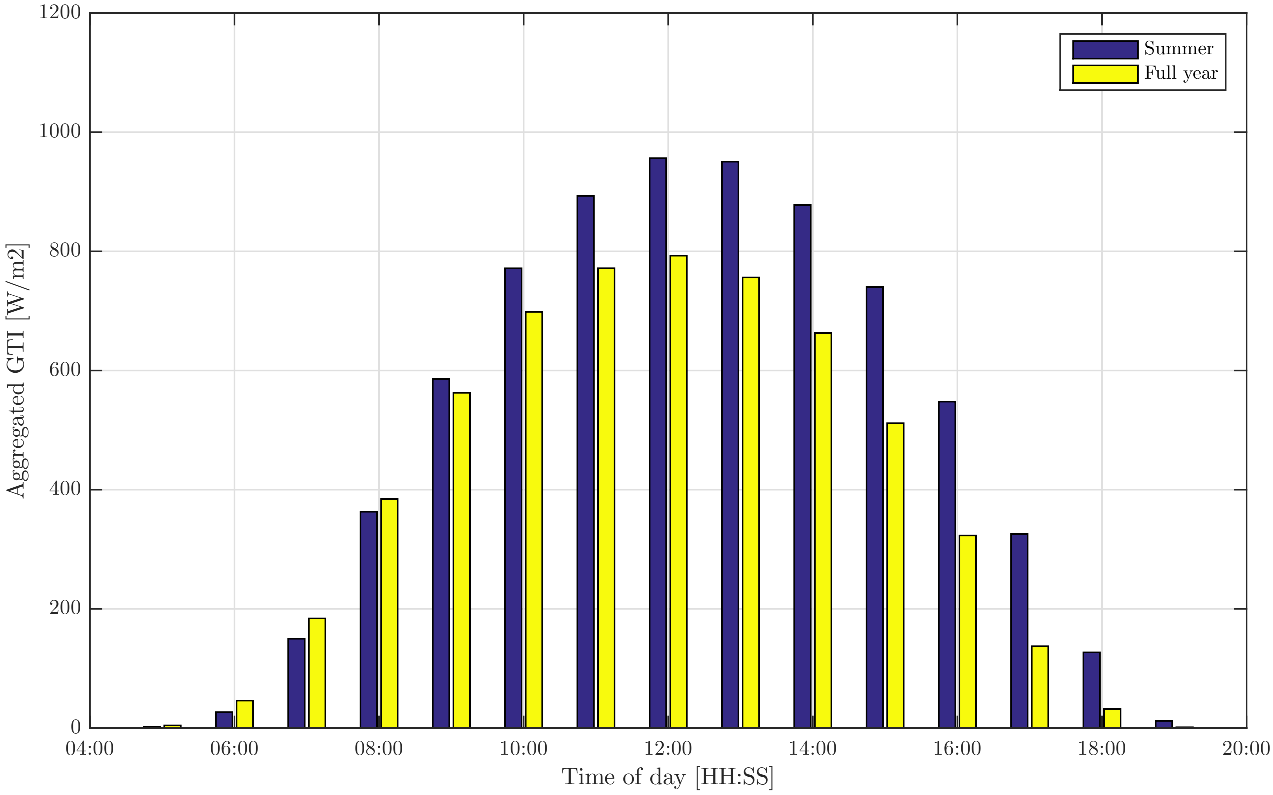

Figure 4.30: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 2 for the combined peak periods during each season.

4.3.3.3 Summary

The power profiles of the optimisation results for this objective function indicate that it may be useful for finding distributions that show reduced midday power peaks while retaining significant levels of one or both of the early morning and late afternoon power components. The full day seems more useful in this regard, as the seasons applicable to the combined peak period consideration are limited by low evening peak energy levels and its results offer no clear advantages. This objective function would likely be much more useful for optimising solar photovoltaic (PV) installations with tracking capabilities, as such systems could produce a much flatter, wider aggregated power profile. The same principle applies to optimisation for an interconnected grid spanning a much larger range in terms of latitude.

Table 4.10 presents the cumulative energy values associated with the optimal distributions found for relevant problem cases analysed with objective function 2, while Table 4.11 presents the variability of the cumulative daily energy associated with each solution. The energy values for the full day results indicate that the diminished midday power levels can result in a significant loss of cumulative annual or seasonal energy relative to the maximum energy available within a selection of locations. Once again, several distributed RSD values prove to be lower than the best RSD for any one location, e.g. the full day, full year cases for site groups 1 and 3.

Table 4.10: Cumulative available energy for problem cases analysed for objective function 2 with GTI profiles.

Problem case description			Cumulative annual energy [MWh/m^2]			Cumulative energy in winter [MWh/m^2]	
Diurnal period	***Site group***	***Seasonal period***	***Full day***	***Morning peak***	***Evening peak***	***Full day***	***Morning peak***
Full day	1	Winter	2236.127	398.842	16.844	503.446	78.293
		Full year	2356.471	386.997	22.048	519.045	71.856
	2	Winter	2332.189	541.027	1.317	584.422	128.312
		Full year	2257.854	505.362	1.945	562.147	115.898
	3	Winter	2382.627	491.619	9.260	570.899	109.446
		Full year	2496.917	408.180	23.553	553.160	77.067
Combined peak periods	1	Summer	2470.998	375.725	27.000	533.889	65.730
		Full year	2101.556	412.087	11.026	486.004	85.490
	2	Summer	2099.824	406.197	5.085	456.543	73.370
		Full year	2332.189	541.027	1.317	584.422	128.312
	3	Summer	2478.384	374.998	27.319	534.846	65.335
		Full year	2141.894	412.755	12.129	493.480	85.003

Table 4.11: RSD of cumulative daily energy for problem cases analysed for objective function 2 with GTI profiles.

Problem case description			RSD of cumulative daily energy [kWh/m^2]	
Diurnal period	***Site group***	***Seasonal period***	***Full year***	***Winter***
Full day	1	Winter	17.4	14.4
		Full year	16.3	12.1
	2	Winter	24.6	17.0
		Full year	26.7	22.0
	3	Winter	18.1	13.3
		Full year	15.4	10.3
Combined peak periods	1	Summer	17.1	11.3
		Full year	21.2	18.1
	2	Summer	29.4	22.1
		Full year	24.6	17.0
	3	Summer	17.1	11.3
		Full year	19.2	16.5

4.3.4 Objective function 3: Maximisation of the daily averaged energy weighted according to Eskom's Megaflex tariff structure

Figures 4.31, 4.33 and 4.35 present the optimal distributions found for site groups 1 to 3 for objective function 3 for the full day, while Figures 4.32, 4.34 and 4.36 present the daily averaged power profiles corresponding to these distributions. Apart from the winter case, which matches the morning peak distribution for objective function 1, the distributions found for site group 1 are almost identical to the results for the full day scenario for objective function 1. The GTI profiles indicate that the morning peak period is prioritised for the winter and autumn cases, while the evening peak period is prioritised for the summer case. Meanwhile, both the spring and full year cases are more balanced throughout the diurnal cycle, with a slight bias towards the evening peak.

Once again, the results for site group 2 match the full day distributions for objective function 1 very closely, with the exception of the spring case, which corresponds to the morning peak distribution for objective function 1. Since the diurnal variation within site group 2 is limited by a lack of longitudinal diversity, the GTI profiles show a similar distribution with the winter profile presenting a slight bias and the spring profile presenting a prominent bias towards the morning peak period.

The results for site group 3 also closely resemble the full day distributions for objective function 1, with the winter case reflecting the morning peak result for objective function 1. The GTI profiles show a strong bias towards the morning peak period for the winter case, with the summer case biased towards the evening peak period, and the remaining profiles more evenly balanced around a midday peak.

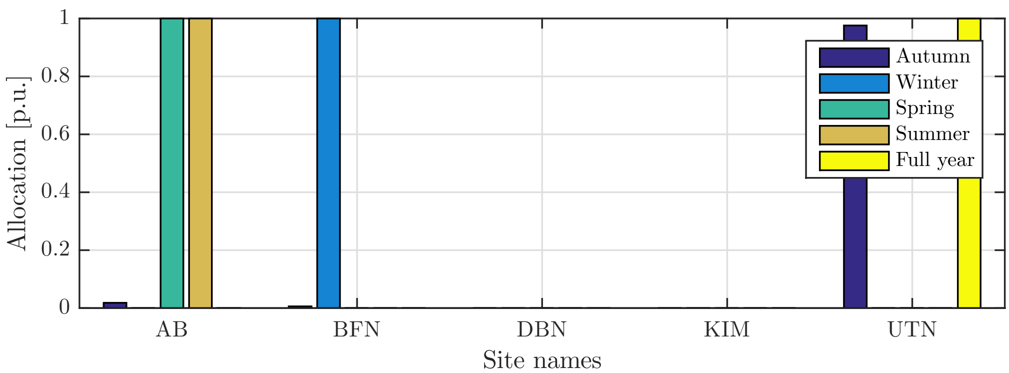

Figure 4.31: Optimal distributions in site group 1 for objective function 3 for the full day during each season.

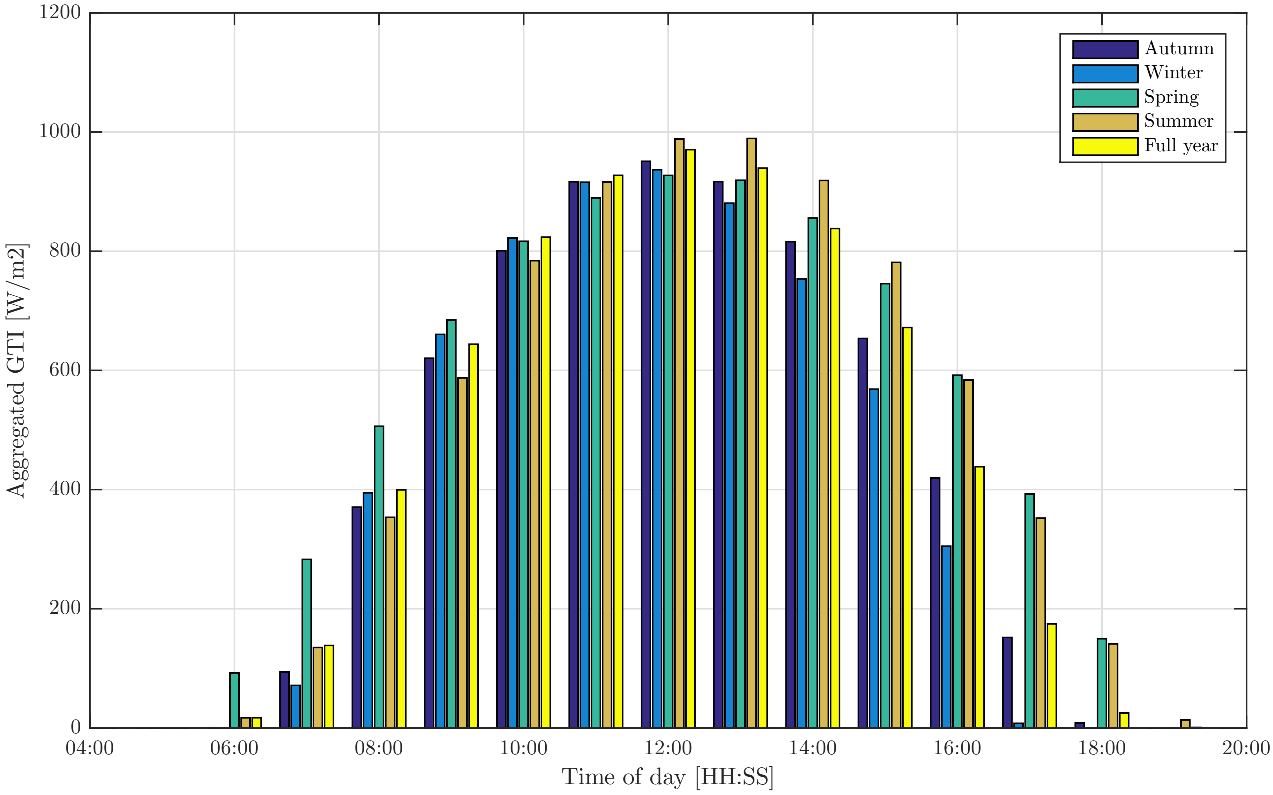

Figure 4.32: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 3 for the full day during each season.

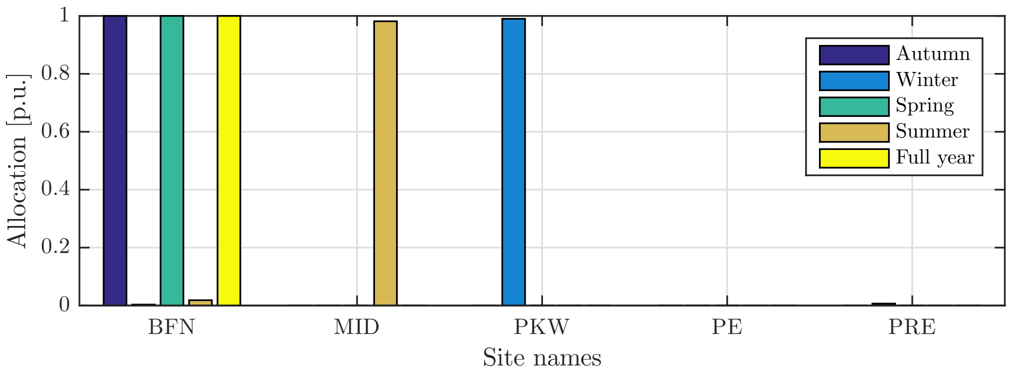

Figure 4.33: Optimal distributions in site group 2 for objective function 3 for the full day during each season.

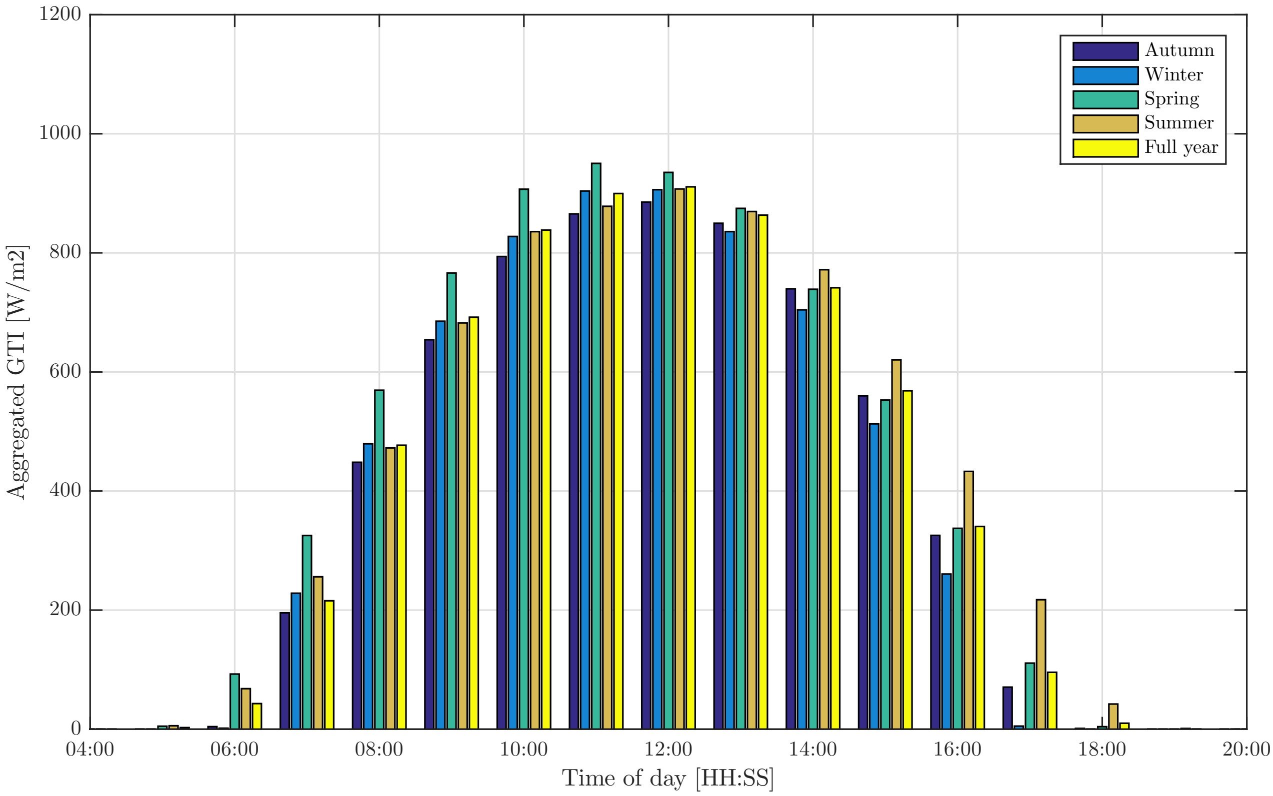

Figure 4.34: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 3 for the full day during each season.

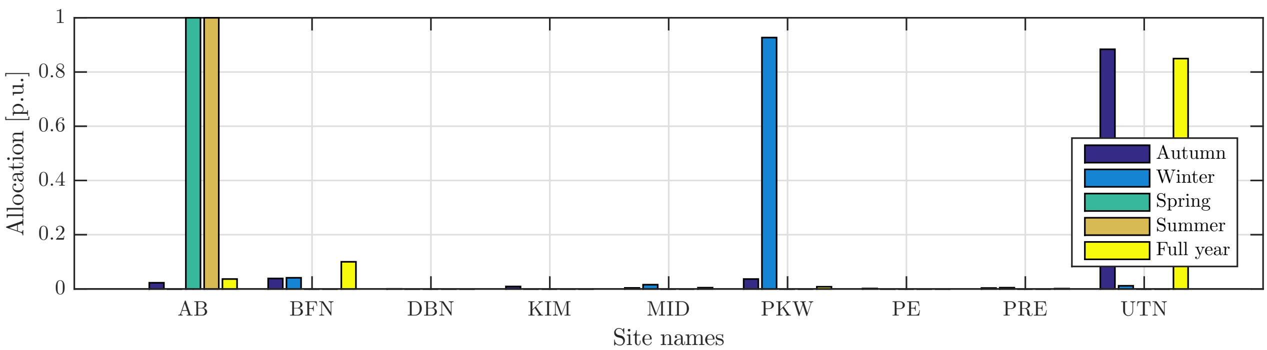

Figure 4.35: Optimal distributions in site group 3 for objective function 3 for the full day during each season.

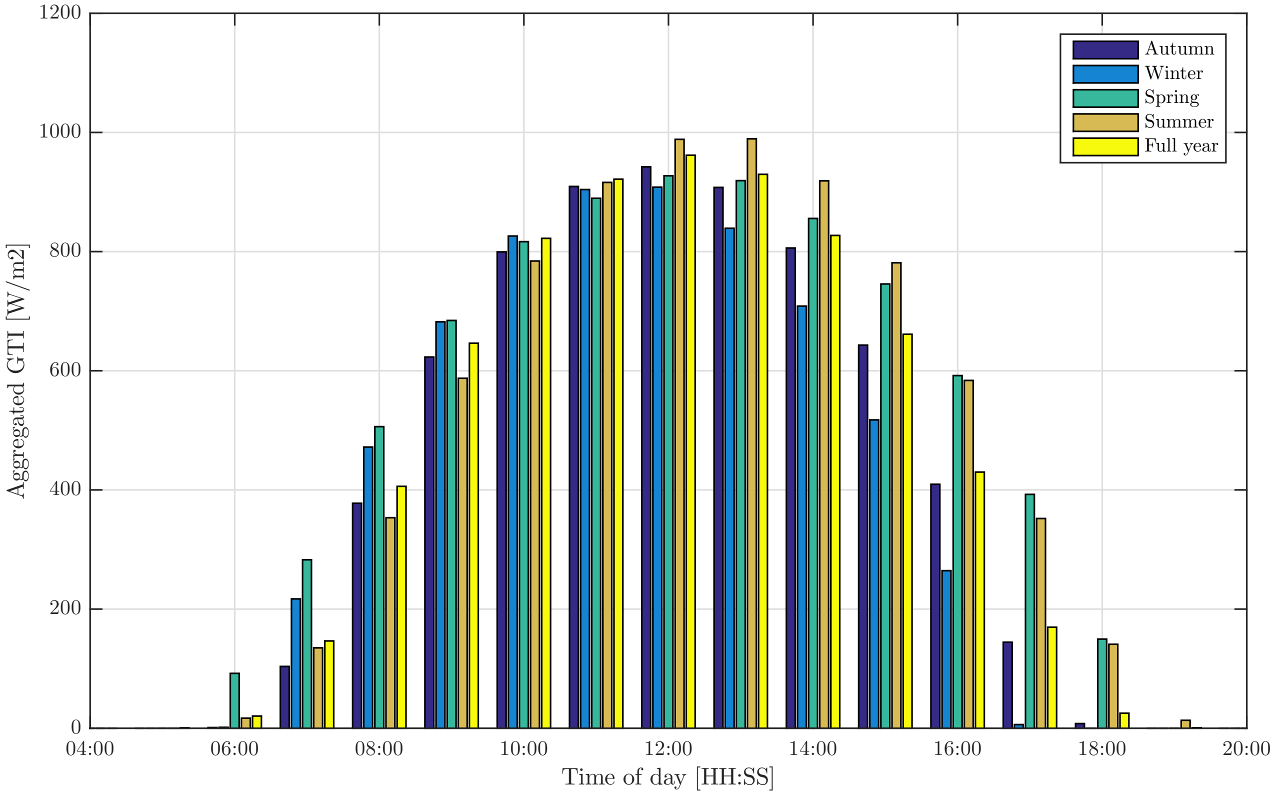

Figure 4.36: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 3 for the full day during each season.

The optimisation results for objective function 3 strongly reflect the full day distributions for objective function 1, which indicates that in many cases simply maximising the daily averaged energy will also deliver the optimal distribution for taking advantage of the Megaflex tariff structure. In certain cases, however, the increased weight of the morning peak period is clearly strong enough to produce a distribution that favours the morning with regard to its power profile. It seems that such distributions occur when prioritising for the morning peak does not significantly reduce the peak power level and there is no option where the evening peak carries noticeable weight.

Table 4.12 presents the cumulative energy values associated with the optimal distributions found for relevant problem cases analysed with objective 3, while Table 4.13 presents the variability of the cumulative daily energy associated with each solution. The cumulative energy values confirm that the winter distributions for site groups 1 and 3 are optimised for maximum morning peak energy rather than maximum cumulative energy (for site group 2 the distribution with the maximum morning peak energy in winter also offers the maximum cumulative energy for the season). The full year distributions for site groups 1 and 2, on the other hand, are optimised for the maximum cumulative annual energy, while the full year distribution for site group 3 exchanged a small portion of the maximum cumulative annual energy for slightly increased morning peak energy.

The variability impact of the aforementioned distributions is varied, depending on the available locations and optimisation focus of each problem case. All results with even a slight degree of distribution, however, show improved RSD values from those associated with the dominant location.

Table 4.12: Cumulative available energy for problem cases analysed for objective function 3 with GTI profiles.

Problem case description		Cumulative annual energy [MWh/m^2]			Cumulative energy in winter [MWh/m^2]	
Site group	***Seasonal period***	***Full day***	***Morning peak***	***Evening peak***	***Full day***	***Morning peak***
1	Winter	2444.946	505.358	3.698	581.101	103.609
	Full year	2558.220	431.325	9.424	592.690	83.581
2	Winter	2332.048	540.664	1.329	584.257	128.145
	Full year	2444.946	505.358	3.698	581.101	103.609
3	Winter	2340.418	537.007	1.569	583.923	126.141
	Full year	2543.811	437.582	9.520	589.531	85.217

Table 4.13: RSD of cumulative daily energy for problem cases analysed for objective function 3 with GTI profiles.

Problem case description		RSD of cumulative daily energy [kWh/m^2]	
Site group	***Seasonal period***	***Full year***	***Winter***
1	Winter	24.2	17.4
	Full year	17.0	8.4
2	Winter	24.3	16.8
	Full year	24.2	17.4
3	Winter	22.8	15.9
	Full year	15.1	8.1

4.3.5 Objective function 4: Minimisation of the SD for the cumulative daily energy

Figures 4.37, 4.39 and 4.41 present the optimal distributions found for site groups 1 to 3 for objective function 4 for the full day, while Figures 4.38, 4.40 and 4.42 present the daily averaged power profiles corresponding to these distributions. The results for all three site groups show a much higher level of geographical dispersion than any distributions found for the previous objective functions. The strongest preference for a single location is evident in the winter cases for site groups 1 and 3 and, to a lesser degree, the summer cases for site group 1 and 3. This results makes sense given the low levels of variability indicated for these locations in Table 4.7. The GTI profiles generally exhibit lower overall power levels than the equivalent profiles for the full day objective function 1 distributions.

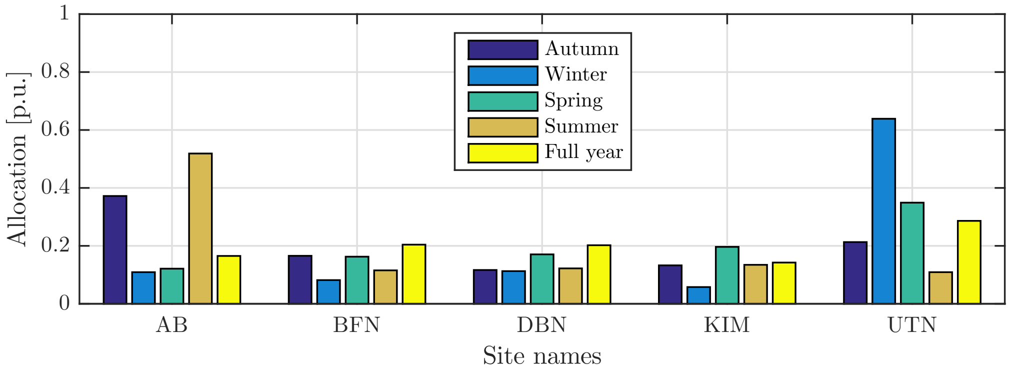

Figure 4.37: Optimal distributions in site group 1 for objective function 4 for the full day during each season.

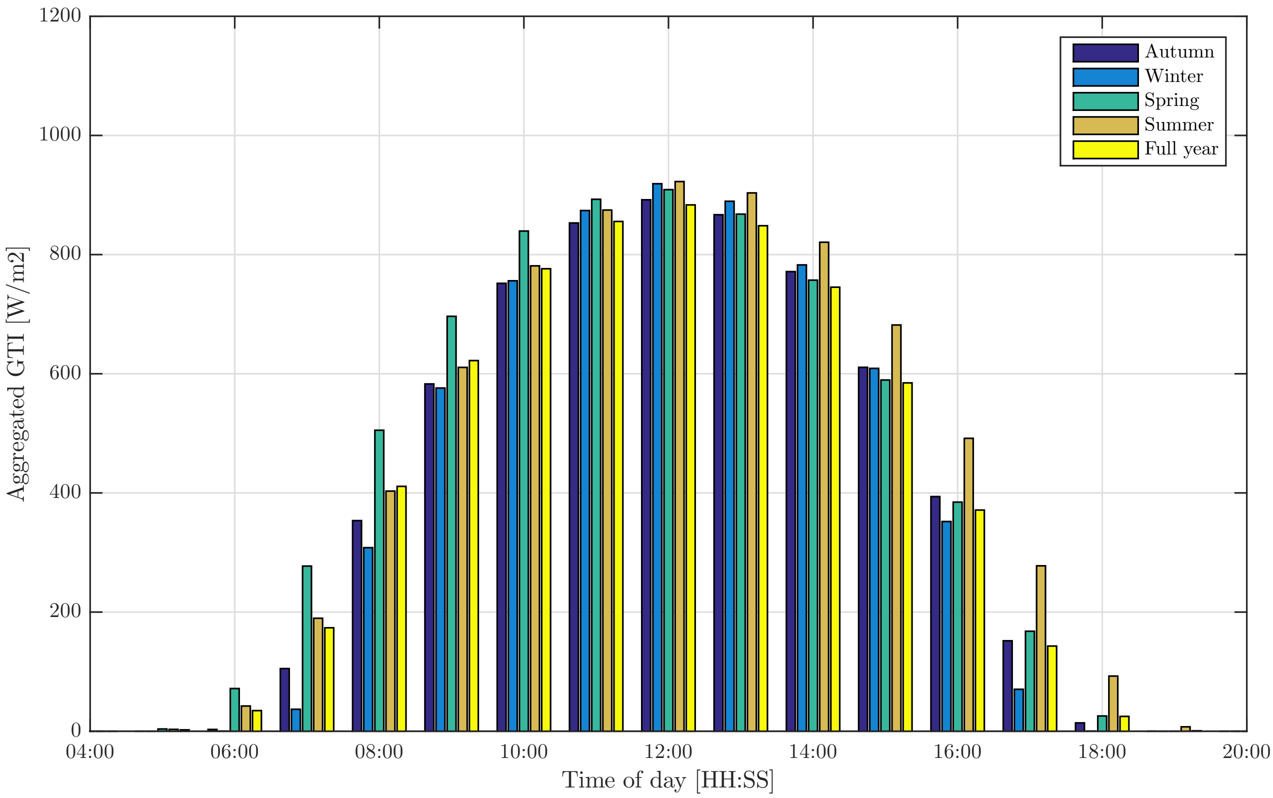

Figure 4.38: Aggregated daily averaged GTI profiles for optimal distributions in site group 1 for objective function 4 for the full day during each season.

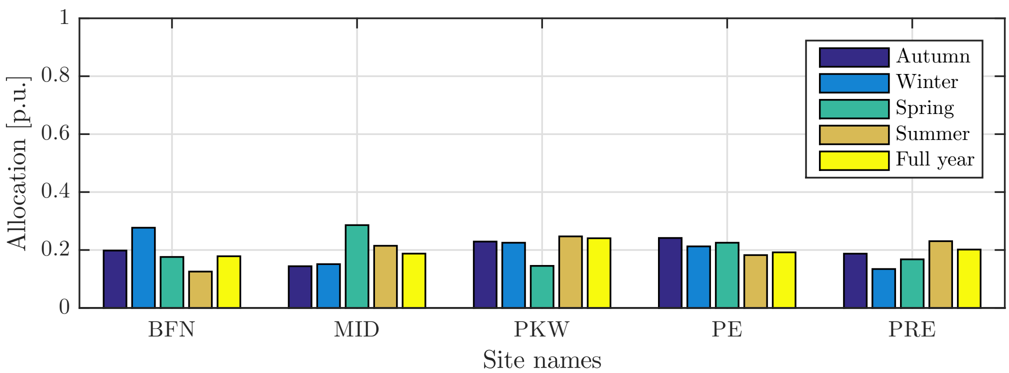

Figure 4.39: Optimal distributions in site group 2 for objective function 4 for the full day during each season.

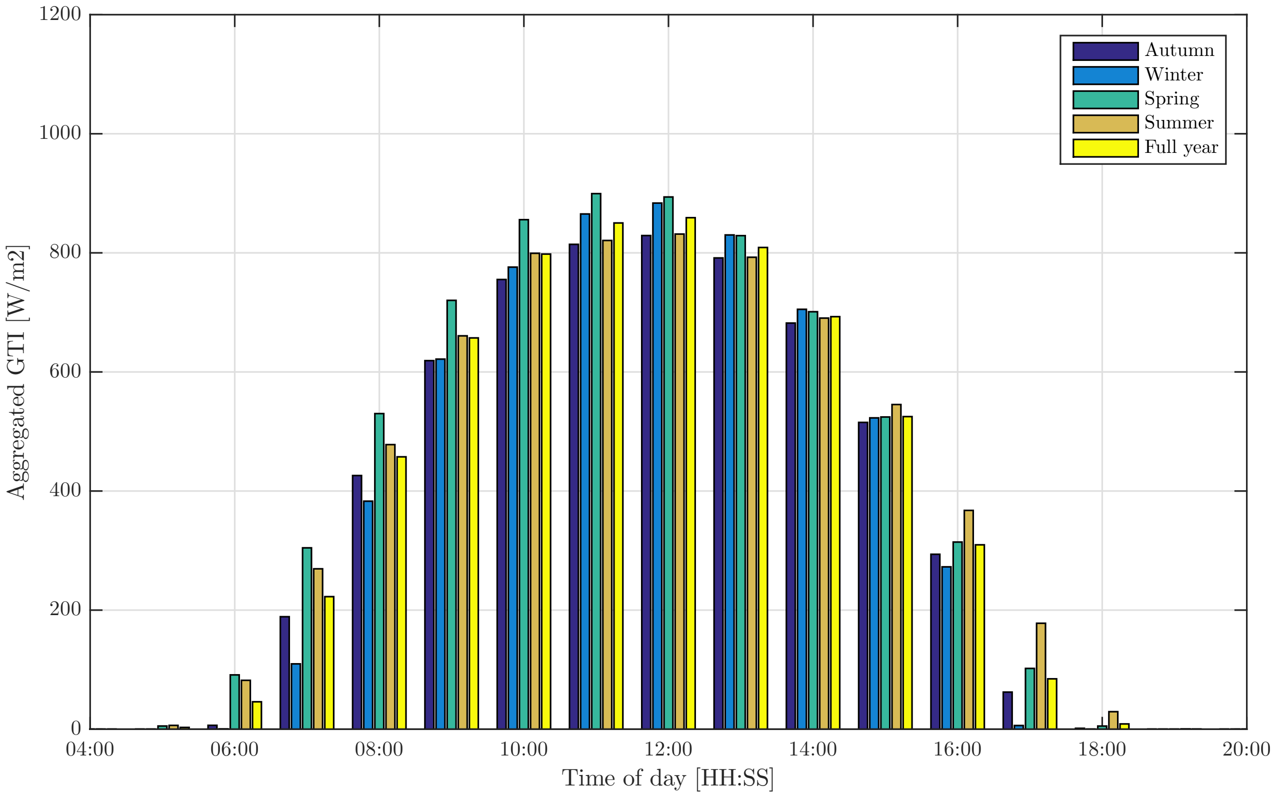

Figure 4.40: Aggregated daily averaged GTI profiles for optimal distributions in site group 2 for objective function 4 for the full day during each season.

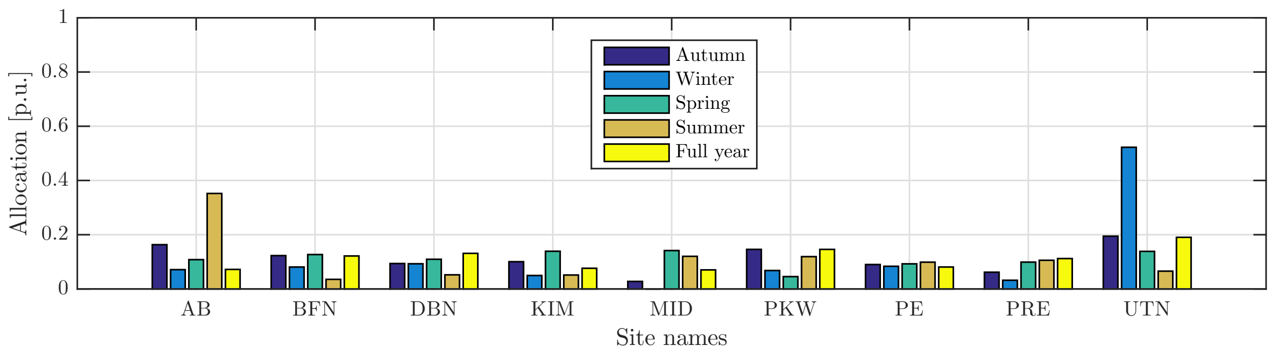

Figure 4.41: Optimal distributions in site group 3 for objective function 4 for the full day during each season.

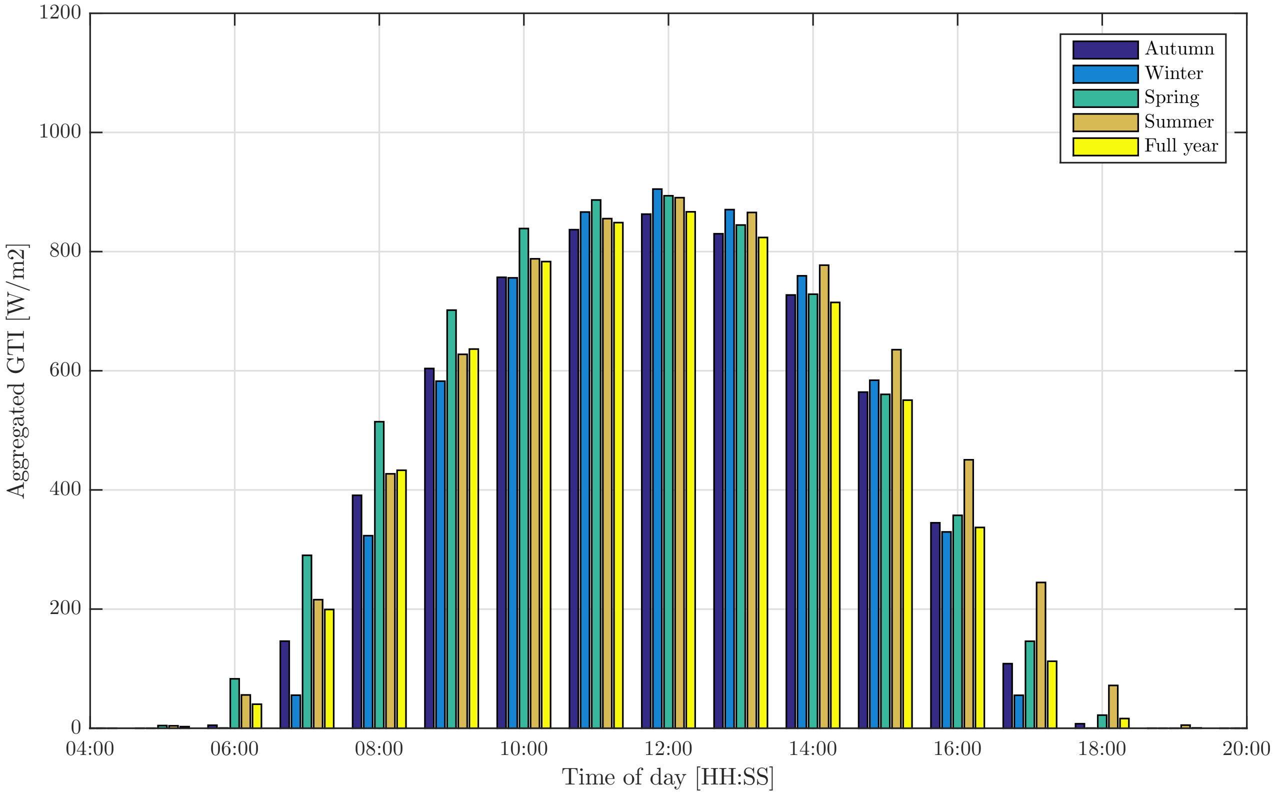

Figure 4.42: Aggregated daily averaged GTI profiles for optimal distributions in site group 3 for objective function 4 for the full day during each season.

Table 4.14 presents the cumulative energy values associated with the optimal distributions found for relevant problem cases analysed with objective function 4, while Table 4.15 presents the variability of the cumulative daily energy associated with each solution. The cumulative energy values show that for each problem case there is a significant loss of cumulative annual energy relative to the maximum available energy in the site group. The RSD values, however, show that the lower energy levels correspond to a notable improvement in the overall variability for each distribution relative to the variability of any single location. The results for site group 3 have the lowest RSD values, indicating that variability continues to be improved by more diverse geographical dispersion (i.e. distribution of the capacity amongst a wider range of locations).

Table 4.14: Cumulative available energy for problem cases analysed for objective function 4 with GTI profiles.

Problem case description		Cumulative annual energy [MWh/m^2]			Cumulative energy in winter [MWh/m^2]	
Site group	***Seasonal period***	***Full day***	***Morning peak***	***Evening peak***	***Full day***	***Morning peak***
1	Winter	2458.445	432.383	10.052	568.011	84.740
	Full year	2364.271	440.484	9.397	547.108	87.728
2	Winter	2316.515	486.842	3.418	549.847	102.508
	Full year	2308.112	488.083	3.283	550.446	103.874
3	Winter	2410.891	442.130	8.308	560.079	88.442
	Full year	2323.580	463.107	6.124	547.953	96.561

Table 4.15: RSD of cumulative daily energy for problem cases analysed for objective function 4 with GTI profiles.

Problem case description		RSD of cumulative daily energy [kWh/m^2]	
Site group	***Seasonal period***	***Full year***	***Winter***
1	Winter	13.3	7.8
	Full year	11.8	9.1
2	Winter	13.1	10.4
	Full year	12.7	10.7
3	Winter	11.8	7.5
	Full year	10.2	8.6

Figures 4.43 and 4.44 and Figures 4.45 and 4.46 represent the relevant probability distributions for the site group 1 winter and full year results, respectively, while Figures 4.47 and 4.48 present the corresponding hourly GTI graphs. These results show that the cumulative daily energy for the optimal winter distribution is limited within 25 % and -20 % of the mean value and goes no lower than 60 % of the maximum value throughout the season, while the cumulative daily energy for the optimal full year distribution is limited within 35 % and -40 % relative to the mean value and goes no lower than 40 % of the maximum value throughout the year. The hourly GTI values for both the winter and full year cases produce a distinct value range throughout the diurnal cycle with relatively few outliers visible.

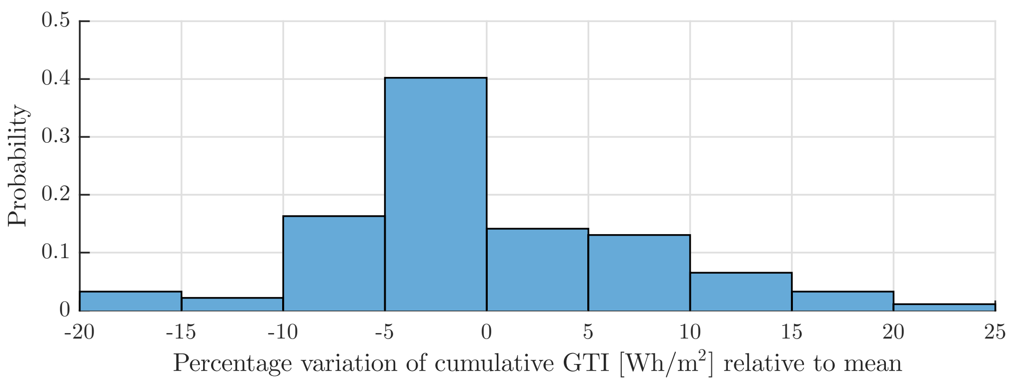

Figure 4.43: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 1 for objective function 4.

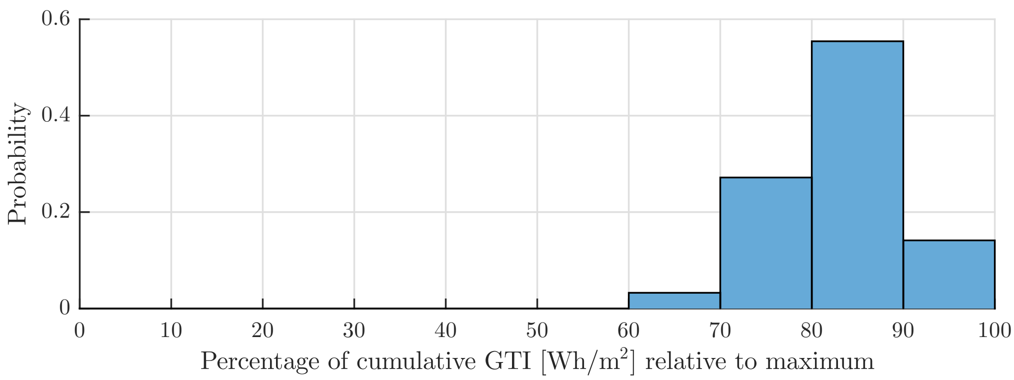

Figure 4.44: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 1 for objective function 4.

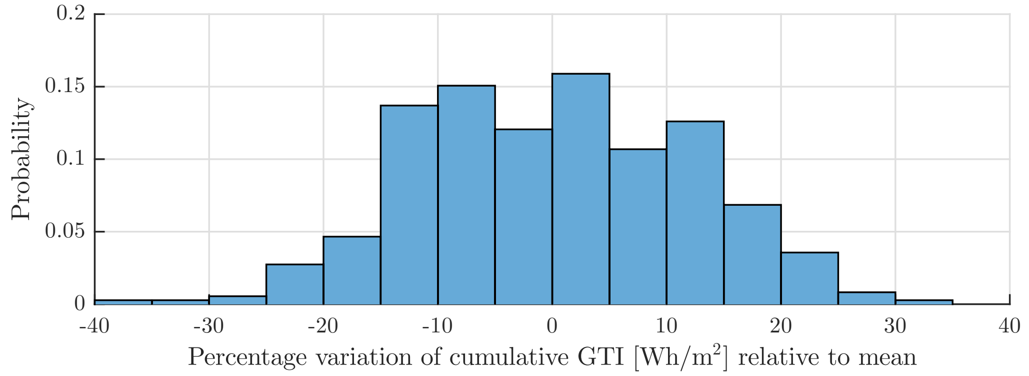

Figure 4.45: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 1 for objective function 4.

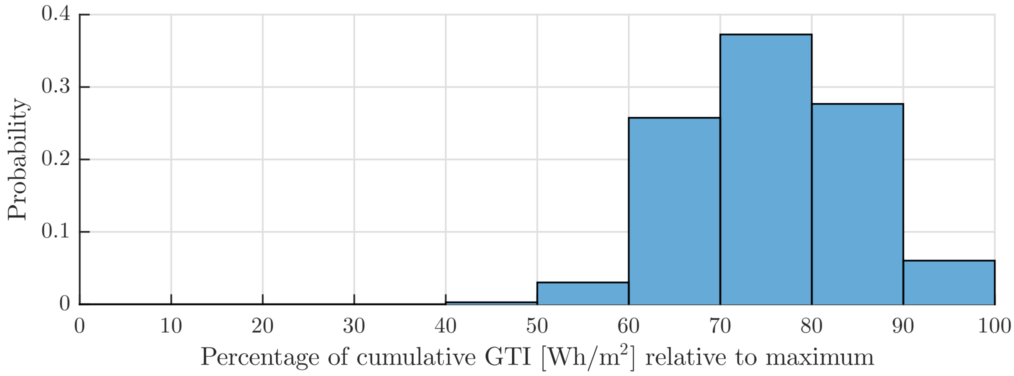

Figure 4.46: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 1 for objective function 4.

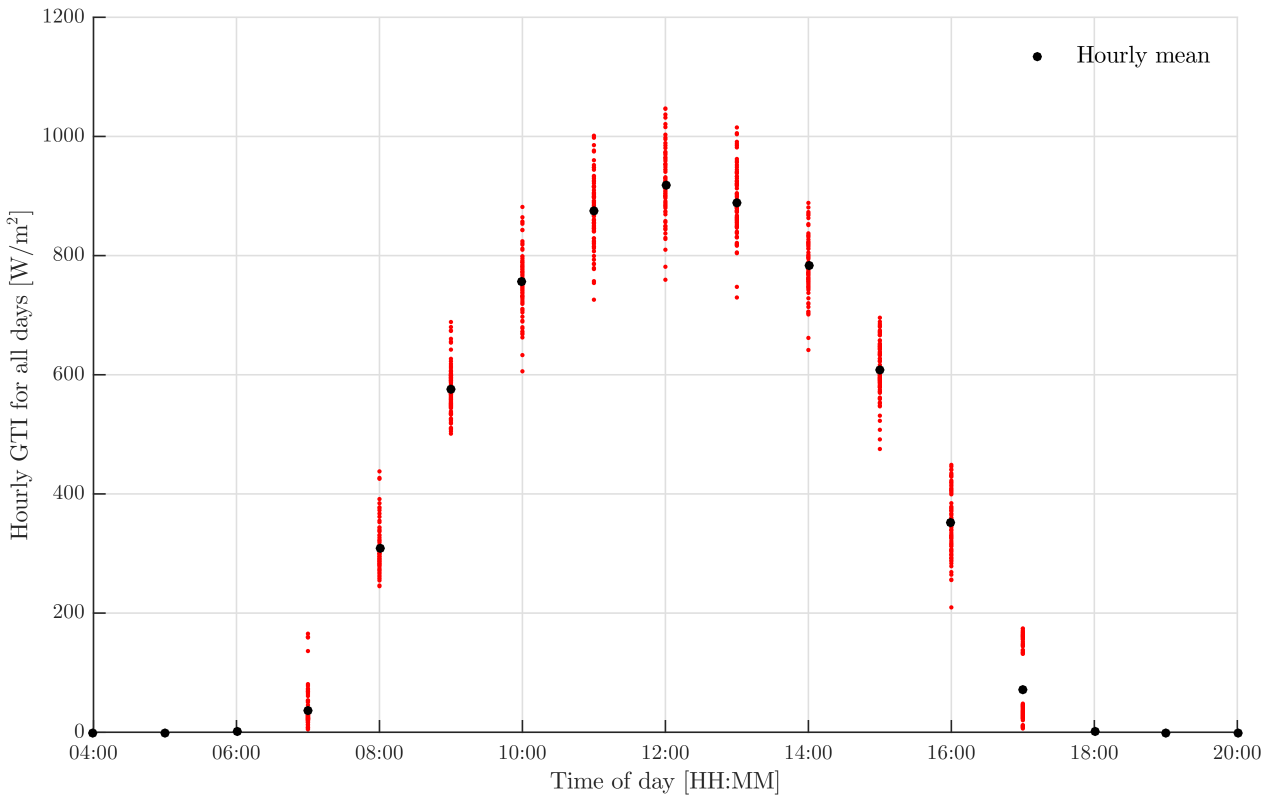

Figure 4.47: Hourly GTI values for the optimal winter distribution in site group 1 for objective function 4.

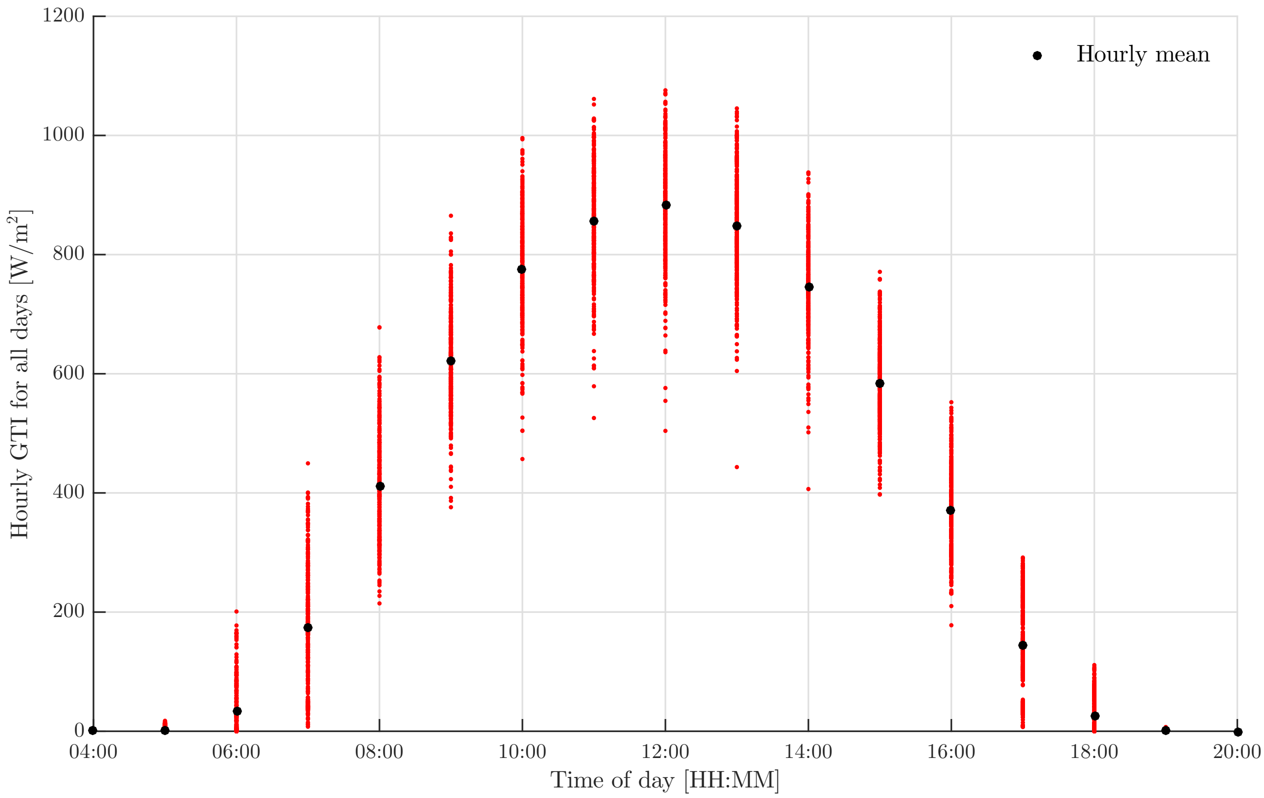

Figure 4.48: Hourly GTI values for the optimal full year distribution in site group 1 for objective function 4.

Figures 4.49 and 4.50 and Figures 4.51 and 4.52 represent the relevant probability distributions for the site group 2 winter and full year results, respectively, while Figures 4.53 and 4.54 present the corresponding hourly GTI graphs. These results reflect the same trends observed for site group 1, with the cumulative daily energy for the optimal winter distribution limited to a variation of 30 % relative to the mean value throughout the season and the cumulative daily energy for the optimal full year distribution limited within of 35 % and -45 % of the mean value throughout the year. Meanwhile the cumulative daily energy also goes no lower than 60 % of the maximum value throughout the season and no lower than 40 % of the maximum value throughout the year. The hourly GTI values for both the winter and full year cases once again produce distinct value ranges throughout the diurnal cycle, although the value range is wider and the outliers more prominent than for the hourly GTI distributions shown for site group 1.

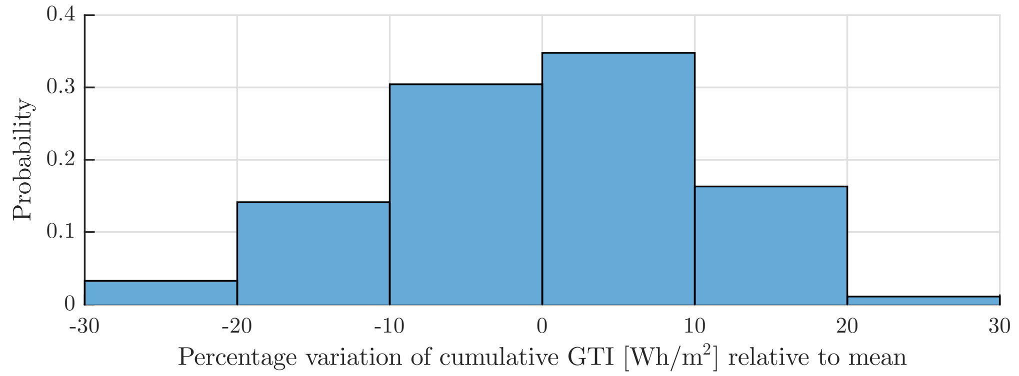

Figure 4.49: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 2 for objective function 4.

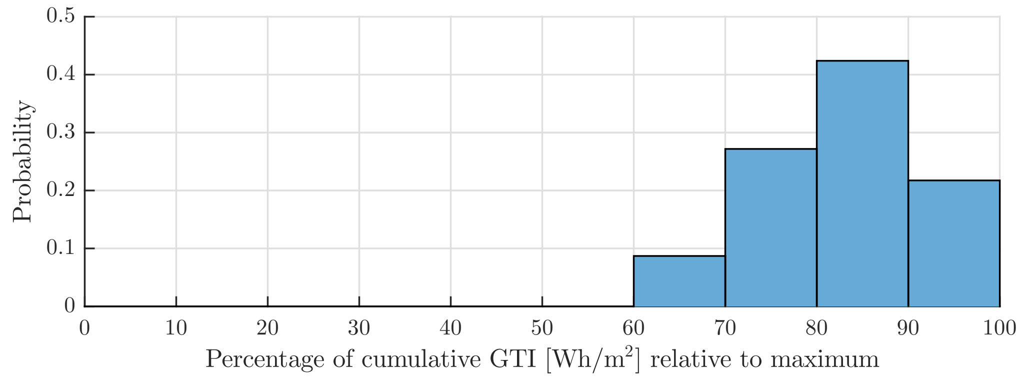

Figure 4.50: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 2 for objective function 4.

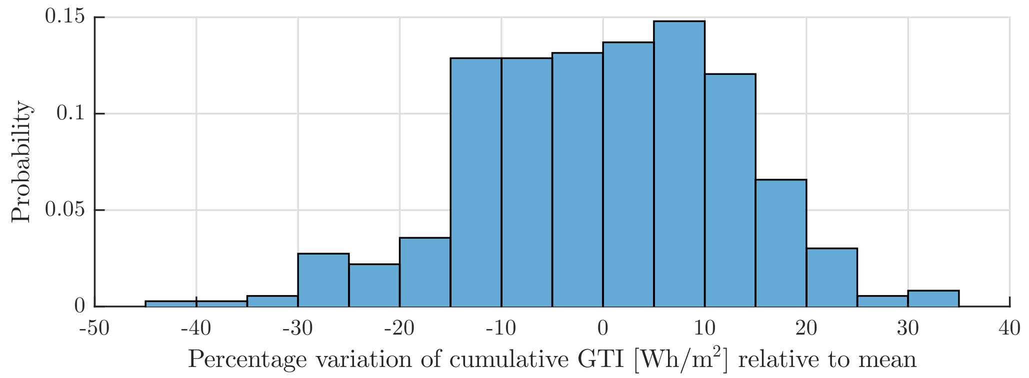

Figure 4.51: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 2 for objective function 4.

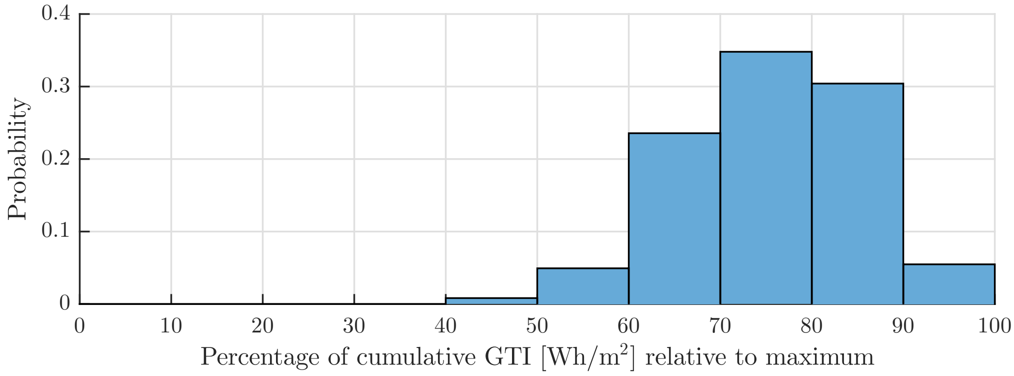

Figure 4.52: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 2 for objective function 4.

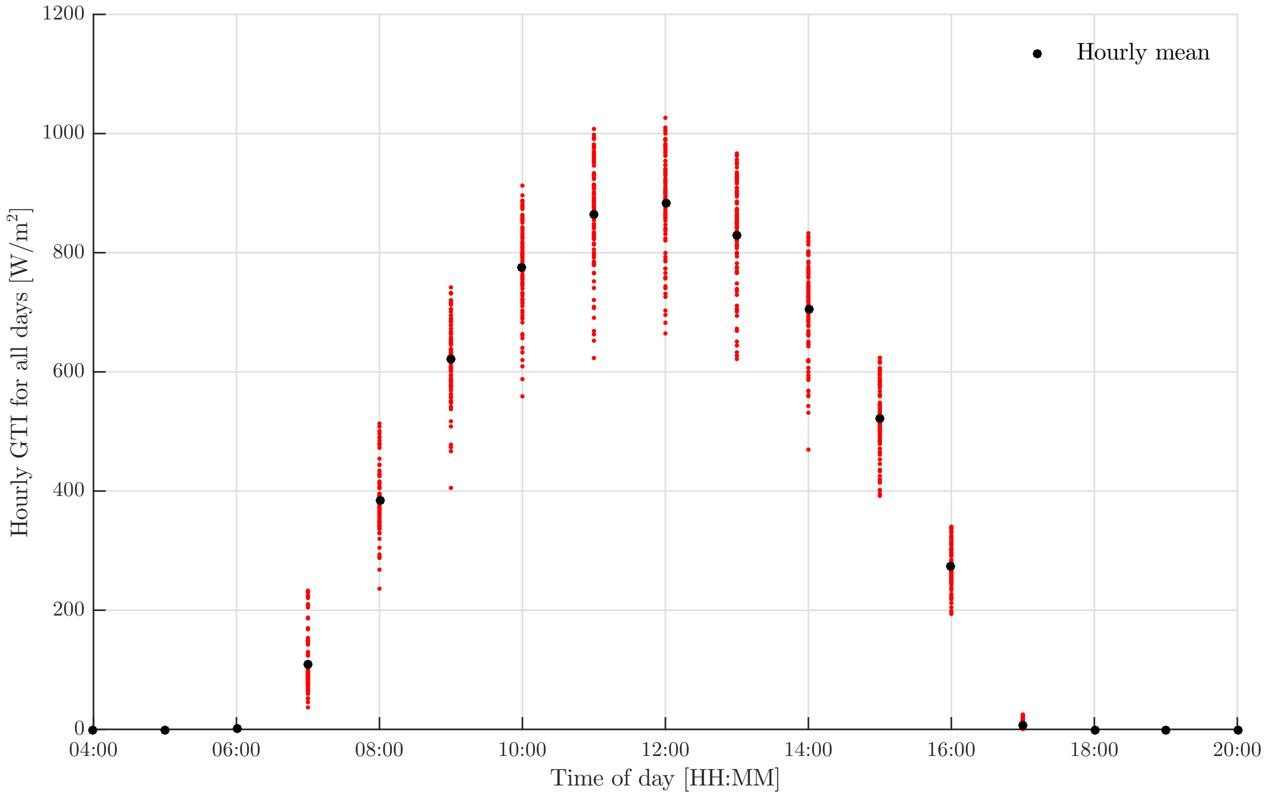

Figure 4.53: Hourly GTI values for the optimal winter distribution in site group 2 for objective function 4.

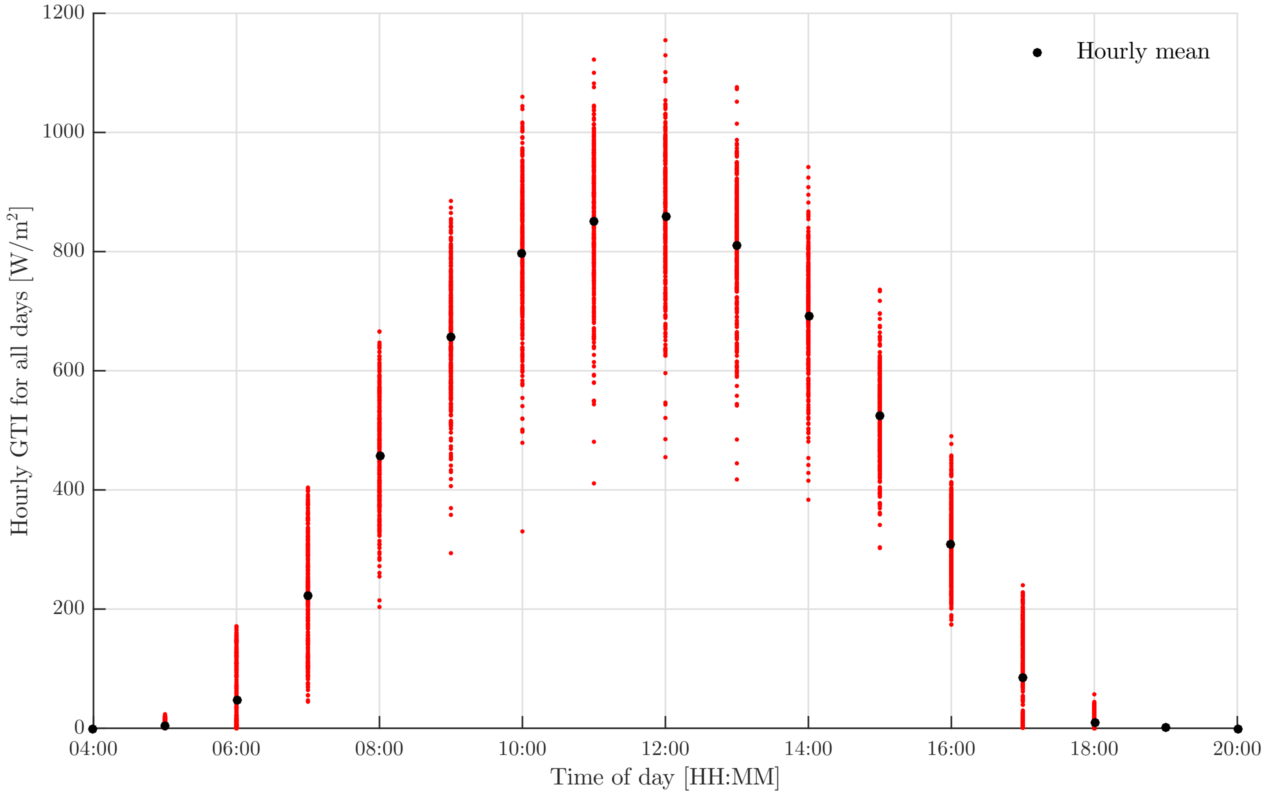

Figure 4.54: Hourly GTI values for the optimal full year distribution in site group 2 for objective function 4.

Figures 4.55 and 4.56 and Figures 4.57 and 4.58 represent the relevant probability distributions for the site group 3 winter and full year results, respectively, while Figures 4.59 and 4.60 present the corresponding hourly GTI graphs. These results show significant improvements compared with the results for both site groups 1 and 2, with the cumulative daily energy for the optimal winter distribution limited to a variation of 20 % relative to the mean value throughout the season and the cumulative daily energy for the optimal full year distribution limited to a variation of 30 % relative to the mean value throughout the year. Meanwhile the cumulative daily energy goes no lower than 70 % of the maximum value throughout the season and no lower than 50 % of the maximum value throughout the year. With regard to the hourly GTI values for both the winter and full year cases, a slight decrease in the frequency and extremity of outliers can be observed relative to the results for site group 1, along with a notable decrease in the value range for the full year case.

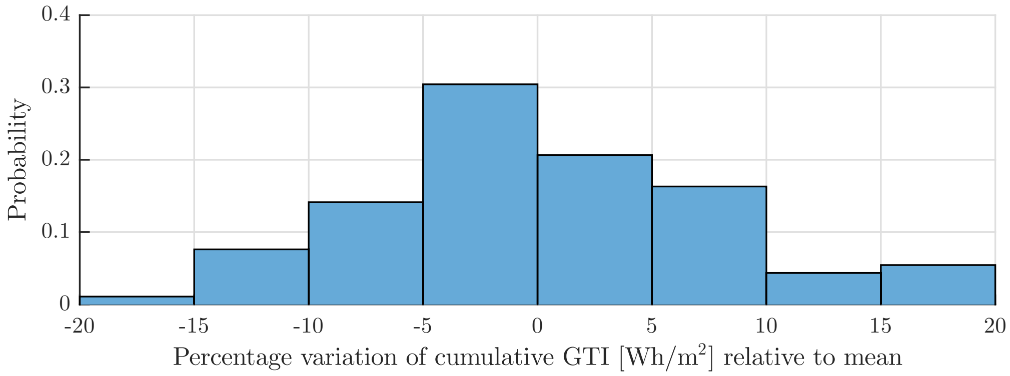

Figure 4.55: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 3 for objective function 4.

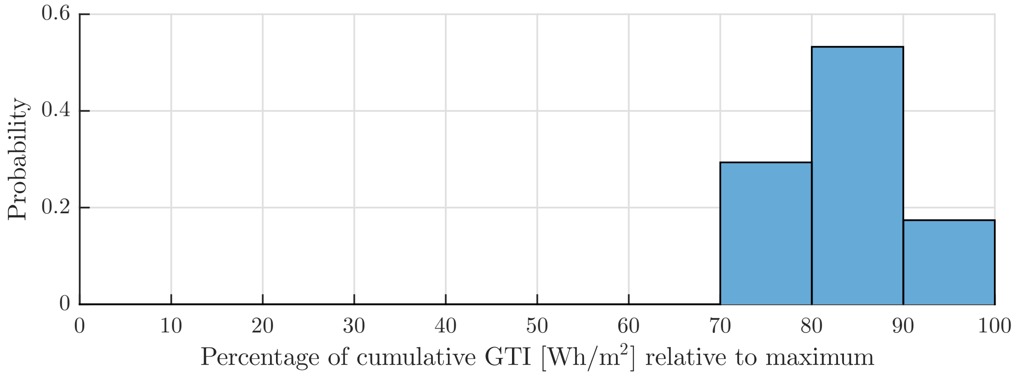

Figure 4.56: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 3 for objective function 4.

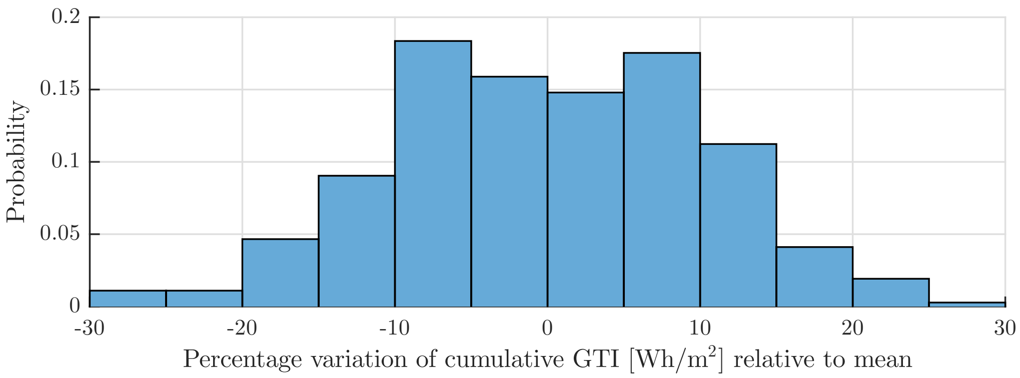

Figure 4.57: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 3 for objective function 4.

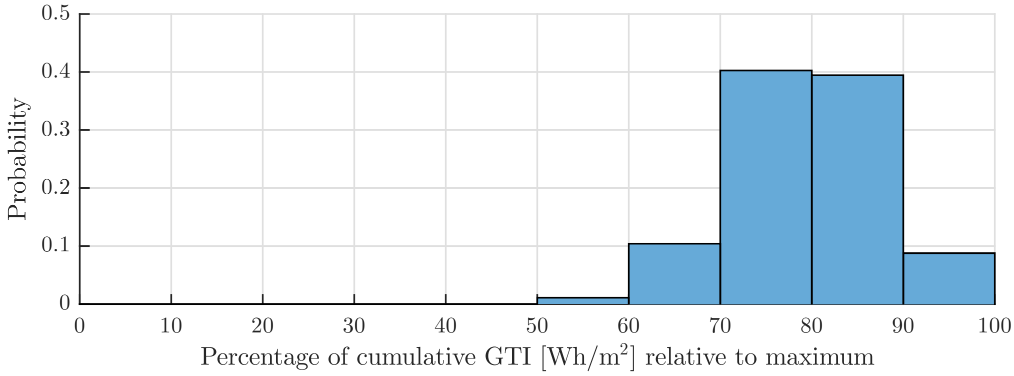

Figure 4.58: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 3 for objective function 4.

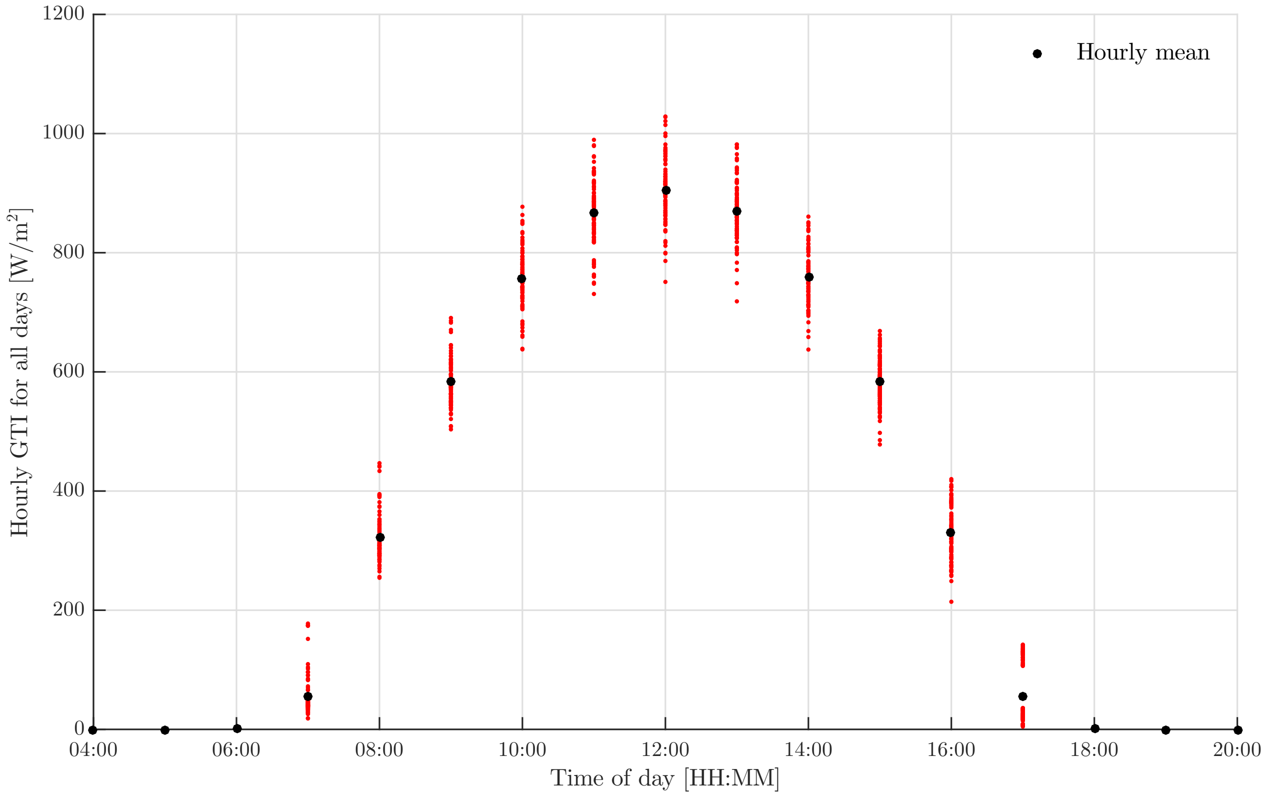

Figure 4.59: Hourly GTI values for the optimal winter distribution in site group 3 for objective function 4.

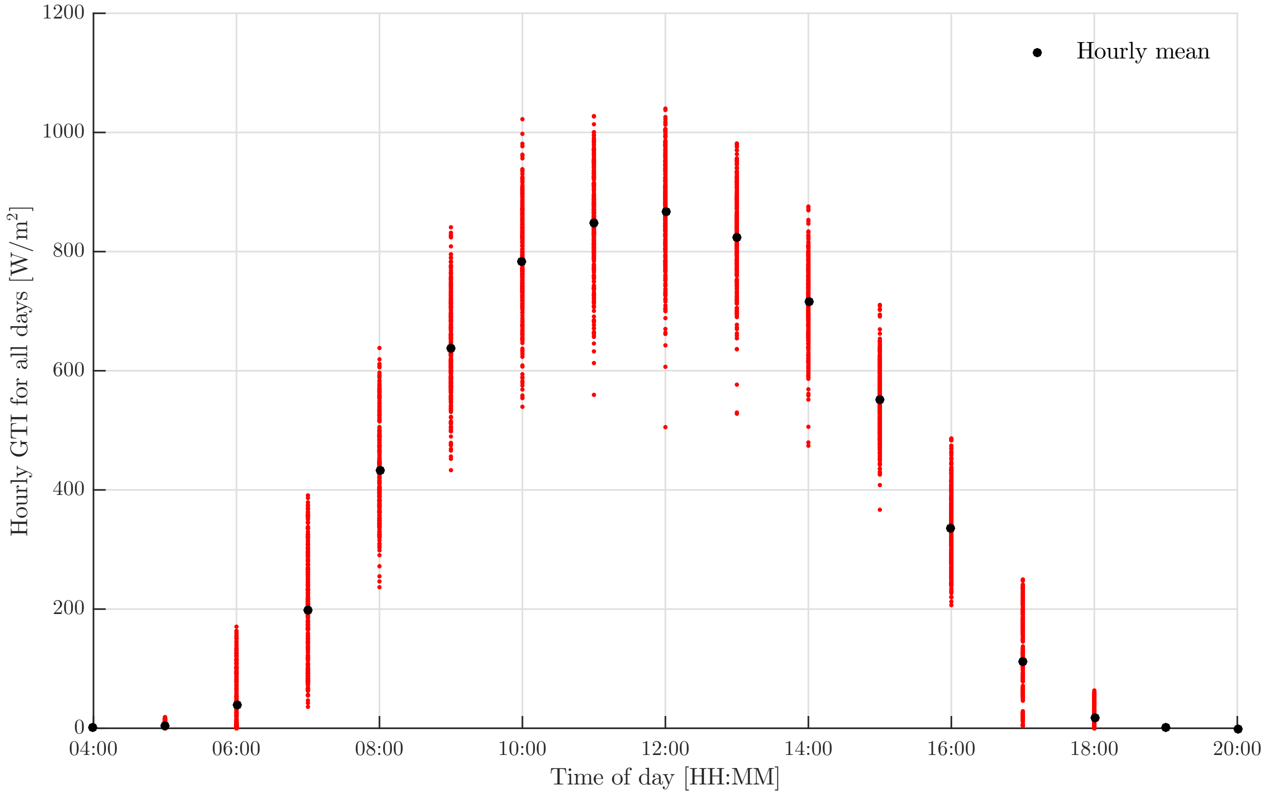

Figure 4.60: Hourly GTI values for the optimal full year distribution in site group 3 for objective function 4.

Overall, the results for the problem cases analysed with objective function 4 indicate that it is a very successful approach for minimising the daily variability within a distribution throughout the duration of a season or the full year. It is also strongly evidenced that the results are improved by increasing the number and geographical diversity of potential locations. As previously indicated, the decrease in variability comes at a cost with regard to the cumulative energy available to the distribution. However, when considering the higher predictability and corresponding mitigation of risk to the grid that is inherent to a distribution with a low level of daily variability, the decrease in cumulative energy could be preferable. After all, the size of solar PV installations within a distribution may be scaled up to produce additional energy, but the effects of high variability is exacerbated for increased generation capacity.

4.3.6 Objective function 5 : Maximisation of negative skewness of the cumulative daily energy

Figures 4.61, 4.63 and 4.65 present the optimal distributions found for site groups 1 to 3 for objective function 5 for the full day, while Figures 4.62, 4.64 and 4.66 present the daily averaged power profiles corresponding to these distributions. The results for all three sites groups show a significant level of dispersion for most problem cases, though the two locations with the highest cumulative energy levels, namely Alexander Bay and Upington, appear to be strongly favoured in site groups 1 and 3. The distributions are therefore less evenly dispersed than those found for the problem cases optimised with objective function 4. The GTI profiles show that these distributions generally produce higher overall energy levels than those found for the equivalent objective function 4 problem cases.

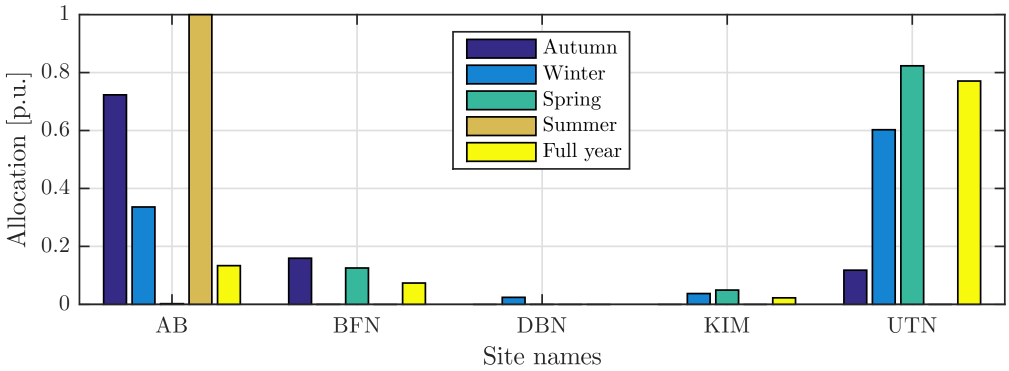

Figure 4.61: Optimal distributions in site group 1 for objective function 5 for the full day during each season.

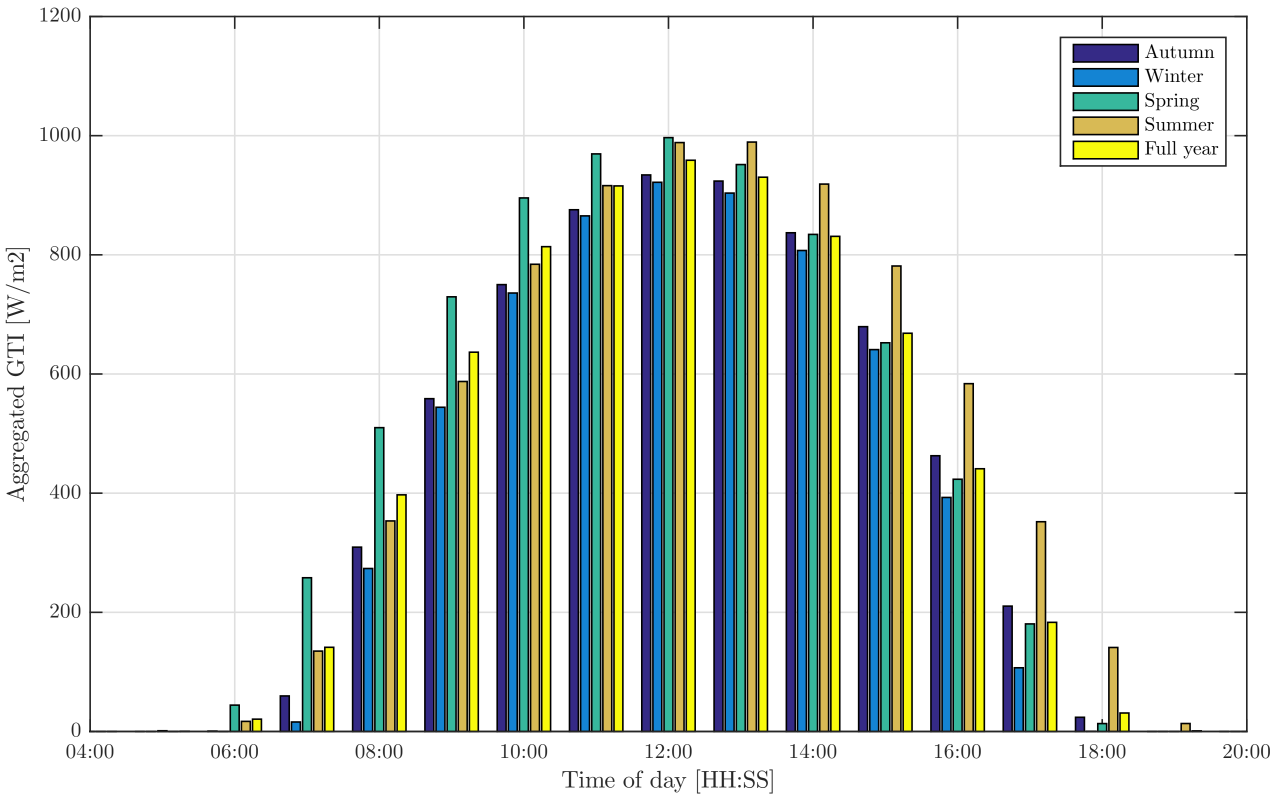

Figure 4.62: Cumulative daily averaged GTI profiles for optimal distributions in site group 1 for objective function 5 for the full day during each season.

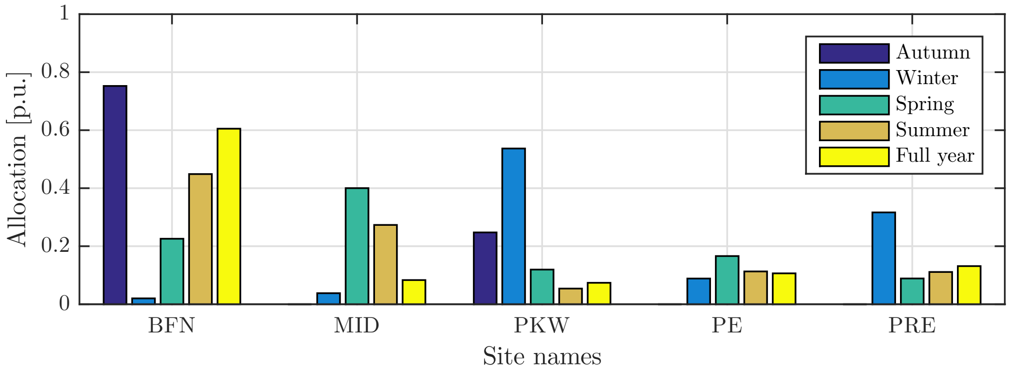

Figure 4.63: Optimal distributions in site group 2 for objective function 5 for the full day during each season.

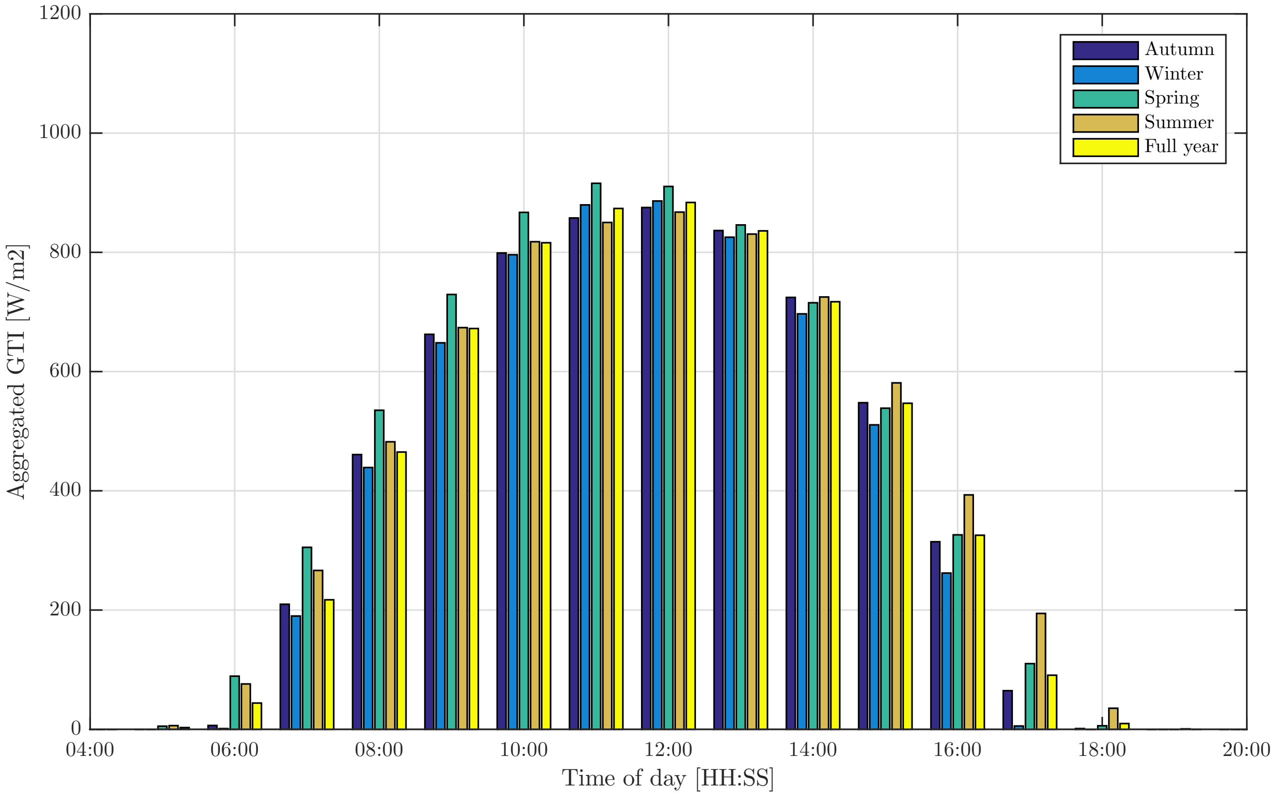

Figure 4.64: Cumulative daily averaged GTI profiles for optimal distributions in site group 2 for objective function 5 for the full day during each season.

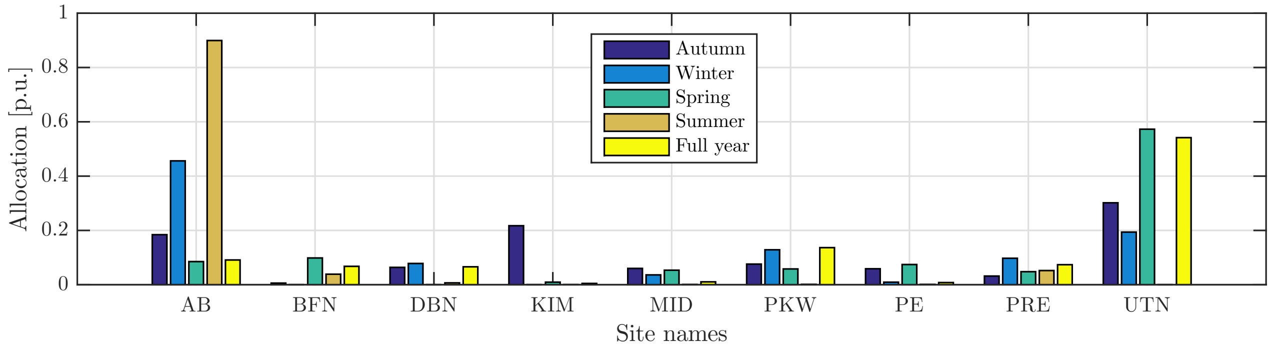

Figure 4.65: Optimal distributions in site group 3 for objective function 5 for the full day during each season.

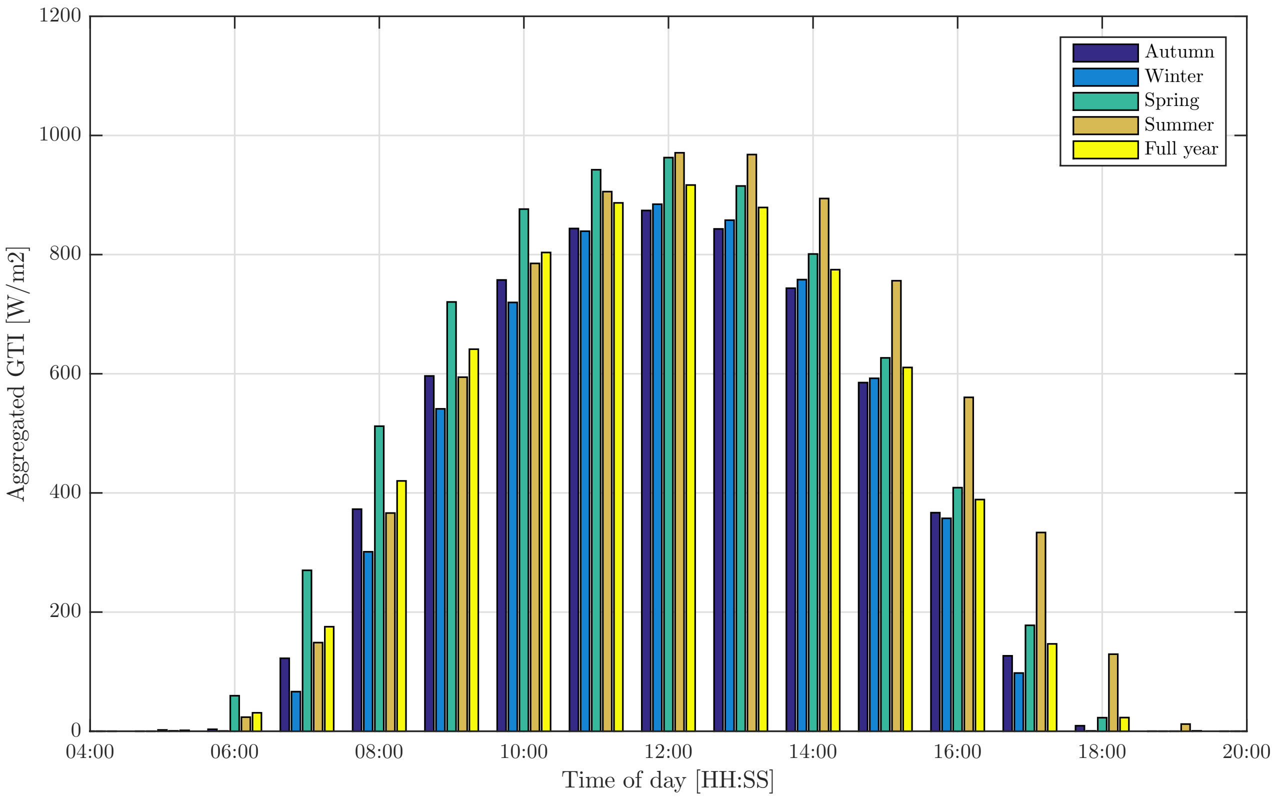

Figure 4.66: Cumulative daily averaged GTI profiles for optimal distributions in site group 3 for objective function 5 for the full day during each season.

Table 4.16 presents the cumulative energy values associated with the optimal distributions found for relevant problem cases analysed with objective function 5, while Table 4.17 presents the variability of the cumulative daily energy associated with each solution. The cumulative energy values confirm that the energy levels attained by the distributions found for objective function 5 are generally higher than that of the equivalent solutions for objective function 4 (though still falling short of the maximum possible energy, to varying extents), with the notable exception of the winter distribution for site group 3. Meanwhile, the RSD values are distinctly higher for each optimised season when compared to the equivalent results for objective function 4. Interestingly, the RSD values produced by the full year distributions in site groups 1 and 2 for only the winter season are lower than their counterparts for the objective function 4 results.

Table 4.16: Cumulative available energy for problem cases analysed for objective function 5 with GTI profiles.

Problem case description		**Cumulative annual energy [MWh/m^2]**			**Cumulative energy in winter [MWh/m^2]**	
Site group	***Seasonal period***	***Full day***	***Morning peak***	***Evening peak***	***Full day***	***Morning peak***
1	Winter	2531.567	411.074	16.132	571.198	76.696
	Full year	2544.152	428.902	11.706	584.197	82.233
2	Winter	2292.874	514.060	2.042	564.924	117.507
	Full year	2372.666	494.367	3.549	563.920	102.770
3	Winter	2434.015	424.959	16.429	553.530	83.616
	Full year	2445.572	451.458	8.653	573.448	92.340

Table 4.17: RSD of cumulative daily energy for problem cases analysed for objective function 5 with GTI profiles.

Problem case description		RSD of cumulative daily energy [kWh/m^2]	
Site group	***Seasonal period***	***Full year***	***Winter***
1	Winter	14.0	8.2
	Full year	14.5	8.0
2	Winter	16.4	13.2
	Full year	16.8	12.3
3	Winter	12.1	9.0
	Full year	11.8	7.6

Figures 4.67 and 4.68 and Figures 4.69 and 4.70 represent the relevant probability distributions for the site group 1 winter and full year results, respectively, while Figures 4.71 and 4.72 present the corresponding hourly GTI graphs. These results show that the cumulative daily energy for the optimal winter distribution is limited within 20 % and -25 % of the mean value and goes no lower than 60 % of the maximum value throughout the season, while the cumulative daily energy for the optimal full year distribution varies between 30 % and -70 % relative to the mean value and upwards of 20 % of the maximum value throughout the year.

When compared with the results for objective function 4, the variability for winter is fairly similar with a slight increase in the negative skewness of the probability distribution relative to the maximum value. This similarity is also reflected in the corresponding hourly GTI profiles. In contrast, the variability of cumulative daily energy for the full year distributions show a pronounced disparity, with the probability distribution relative to the maximum clearly exhibiting increased negative skewness for objective function 5. The adjustment of skewness, however, comes at the cost of a much greater maximum deviation from both the mean and maximum values throughout the year. This behaviour is reflected in the corresponding hourly GTI profile, which shows increased peak power levels with outliers occurring at a significantly higher frequencies and extremities.

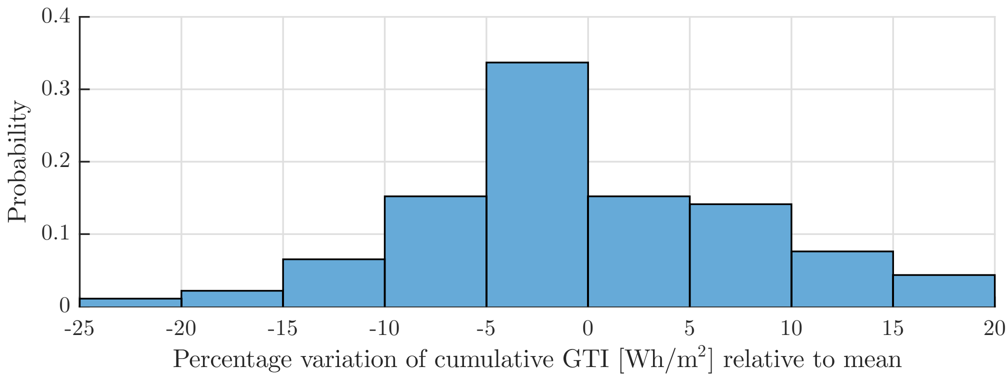

Figure 4.67: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 1 for objective function 5.

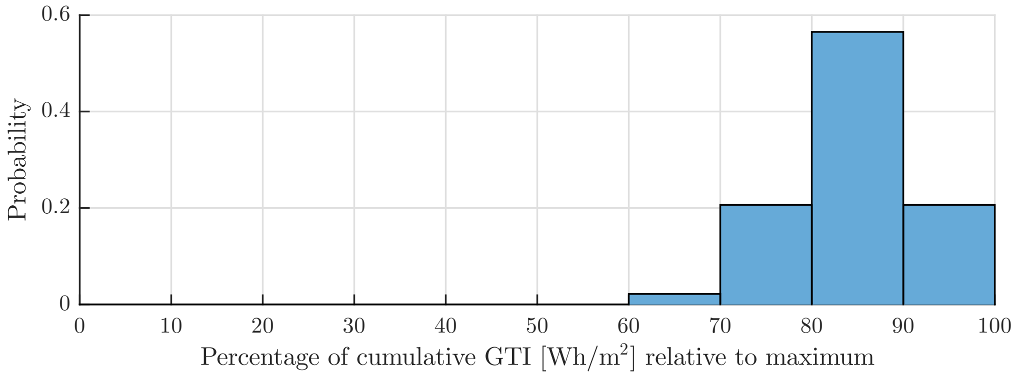

Figure 4.68: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 1 for objective function 5.

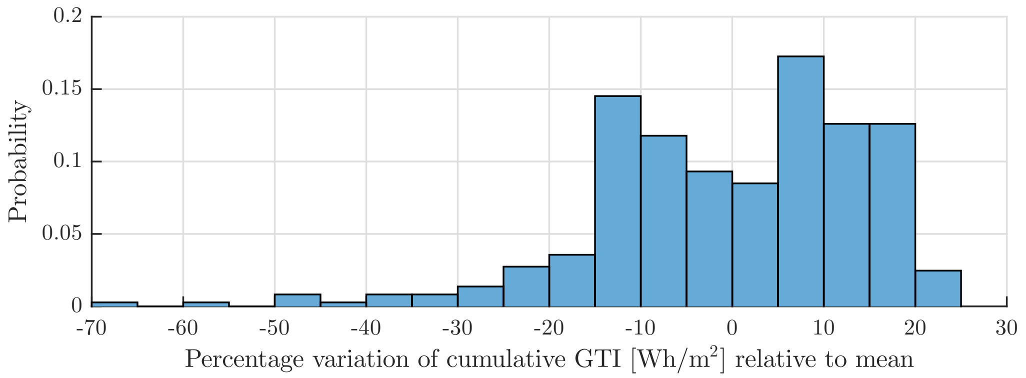

Figure 4.69: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 1 for objective function 5.

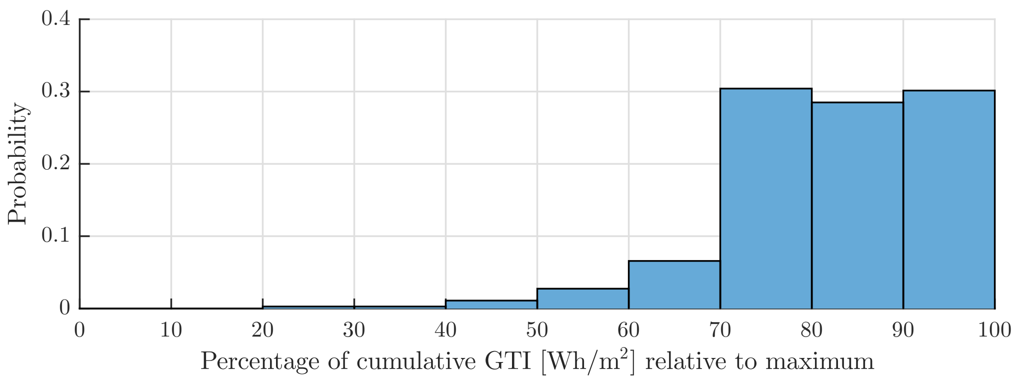

Figure 4.70: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 1 for objective function 5.

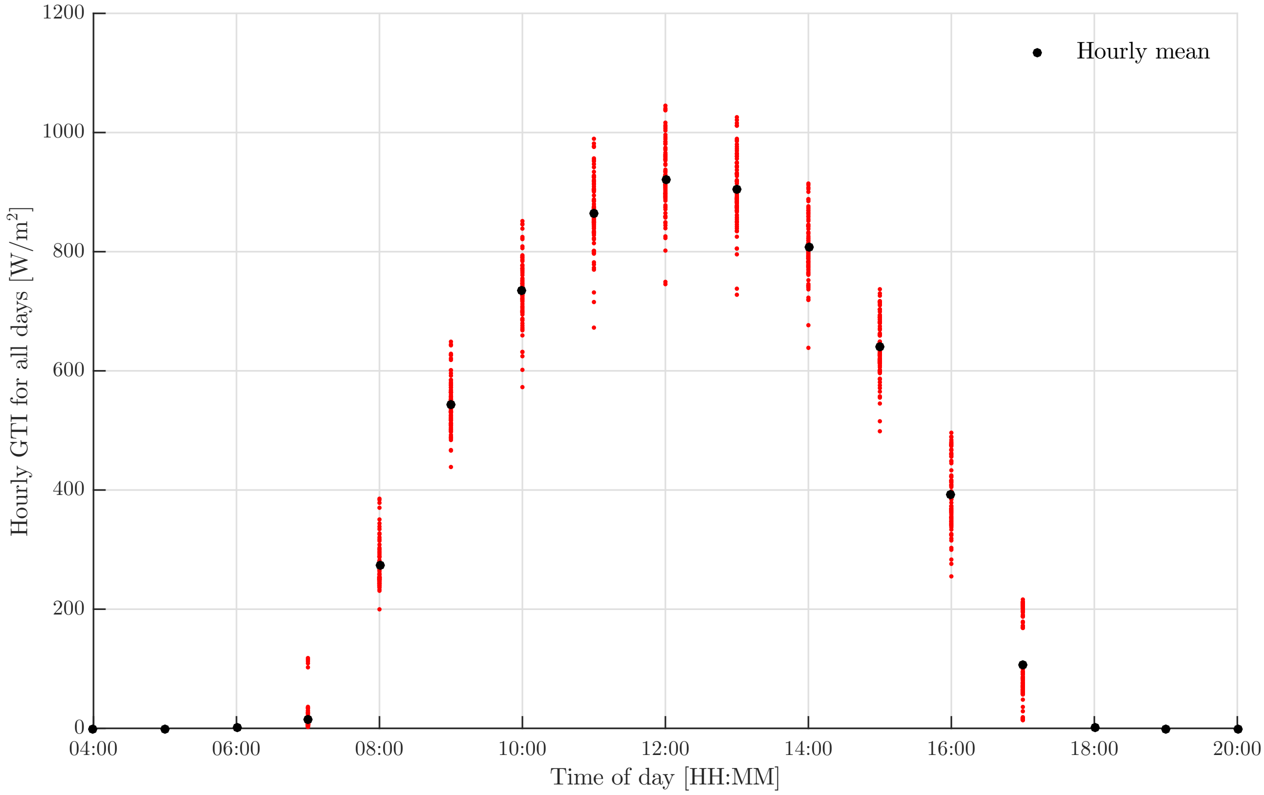

Figure 4.71: Hourly GTI values for the optimal winter distribution in site group 1 for objective function 5.

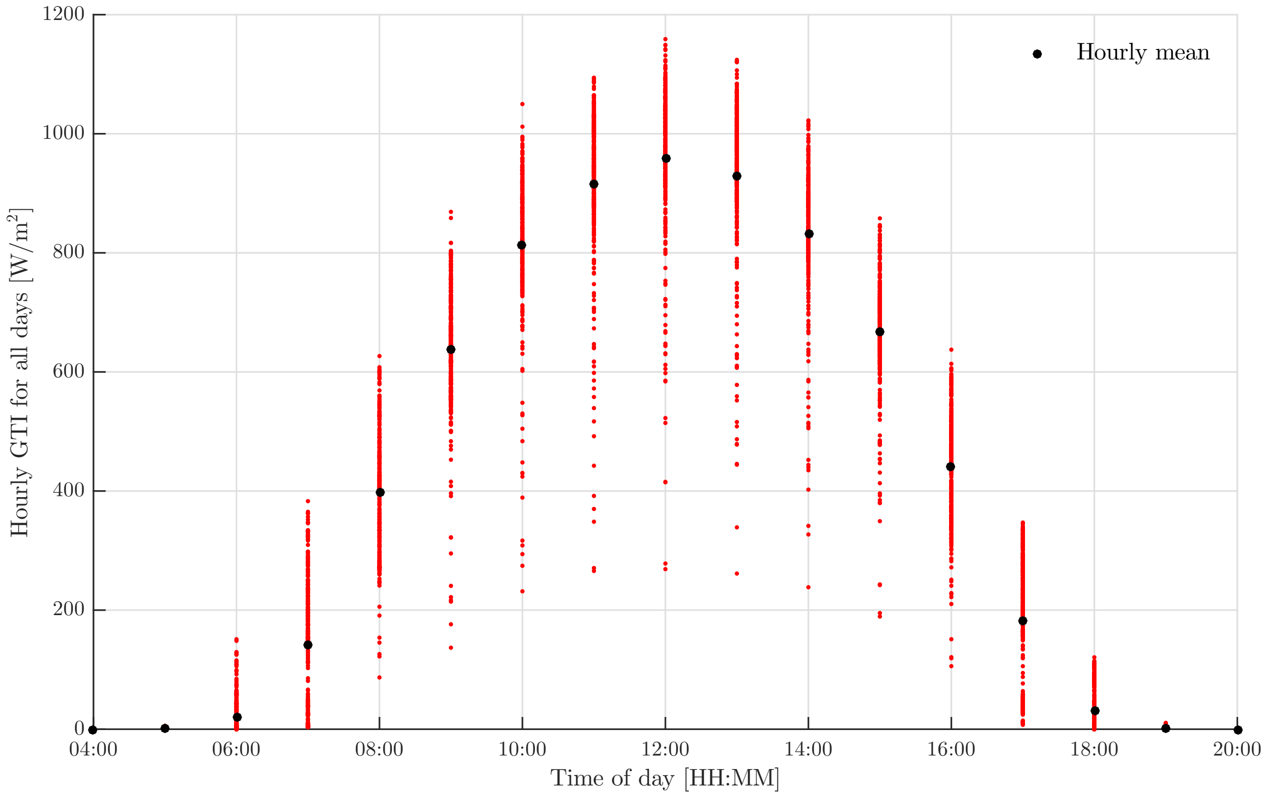

Figure 4.72: Hourly GTI values for the optimal full year distribution in site group 1 for objective function 5.

Figures 4.73 and 4.74 and Figures 4.75 and 4.76 represent the relevant probability distributions for the site group 2 winter and full year results, respectively, while Figures 4.77 and 4.78 present the corresponding hourly GTI graphs. These results show that the cumulative daily energy for the optimal winter distribution varies within -40 % and 20 % of the mean value and goes no lower than 50 % of the maximum value throughout the season, while the cumulative daily energy for the optimal full year distribution varies between 30 % and -70 % relative to the mean value and upwards of 30 % of the maximum value throughout the year.

When compared with the results for objective function 4, the variability of cumulative daily energy for both the winter and full year distributions show a distinct increase in negative skewness of the probability distribution relative to the mean. Once again, this comes at the cost of greater maximum deviation relative to both the mean and the maximum value, which is reflected in the increased downwards dispersion of values in the hourly GTI profiles for both the winter and full year cases.

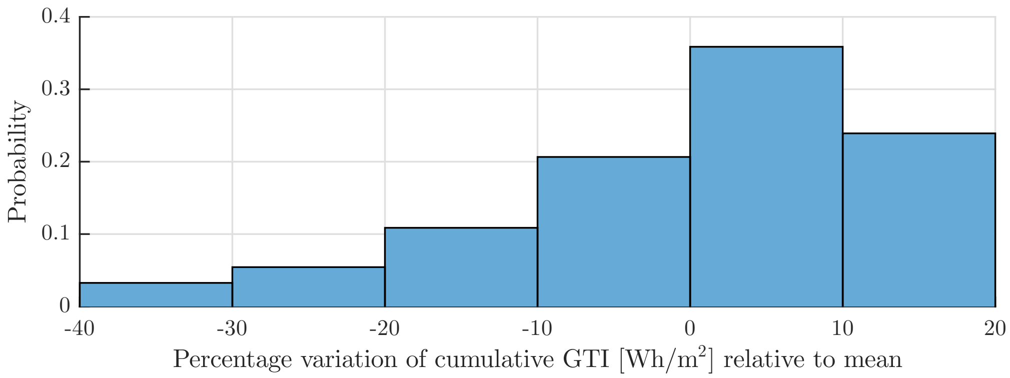

Figure 4.73: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 2 for objective function 5.

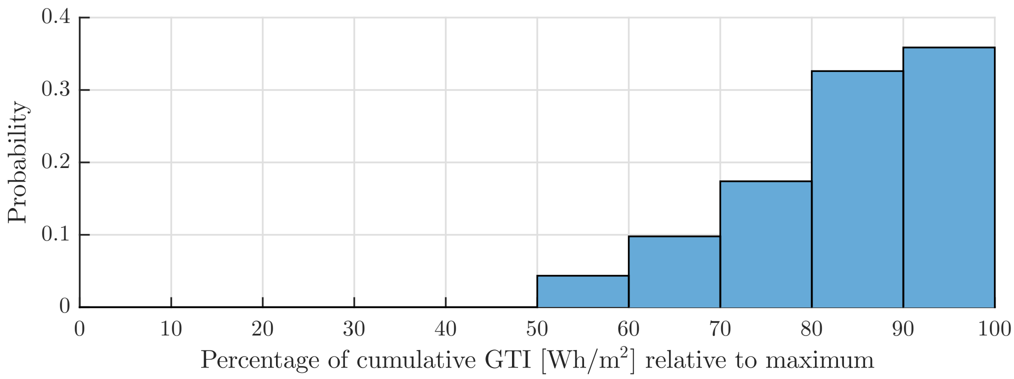

Figure 4.74: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 2 for objective function 5.

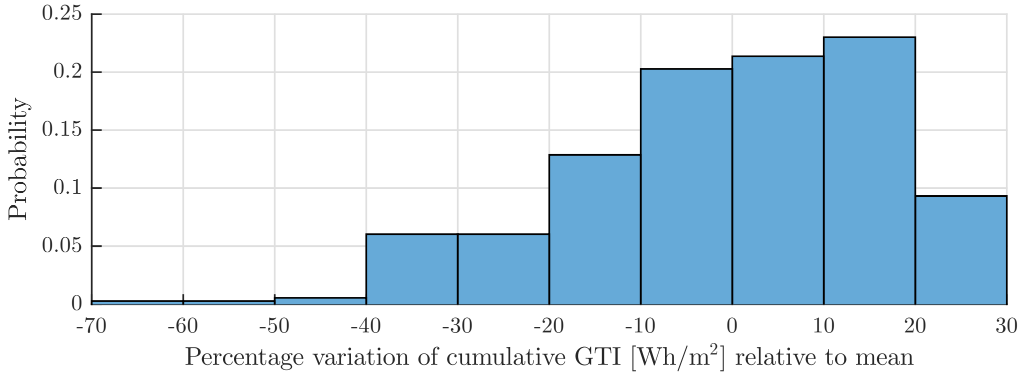

Figure 4.75: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 2 for objective function 5.

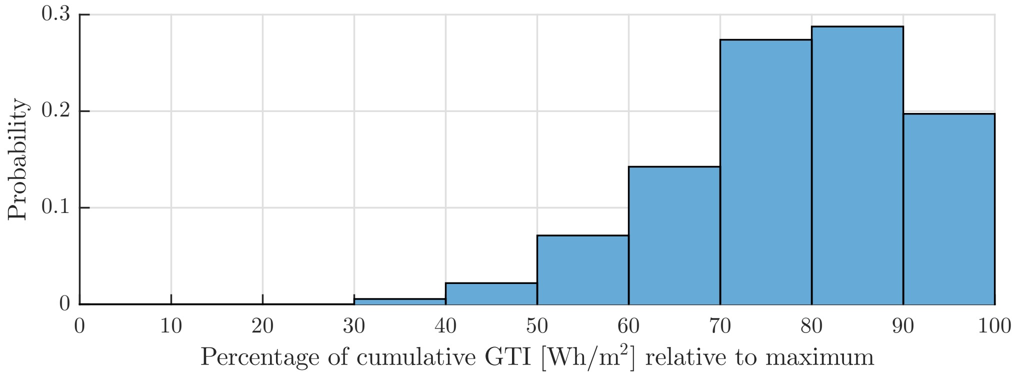

Figure 4.76: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 2 for objective function 5.

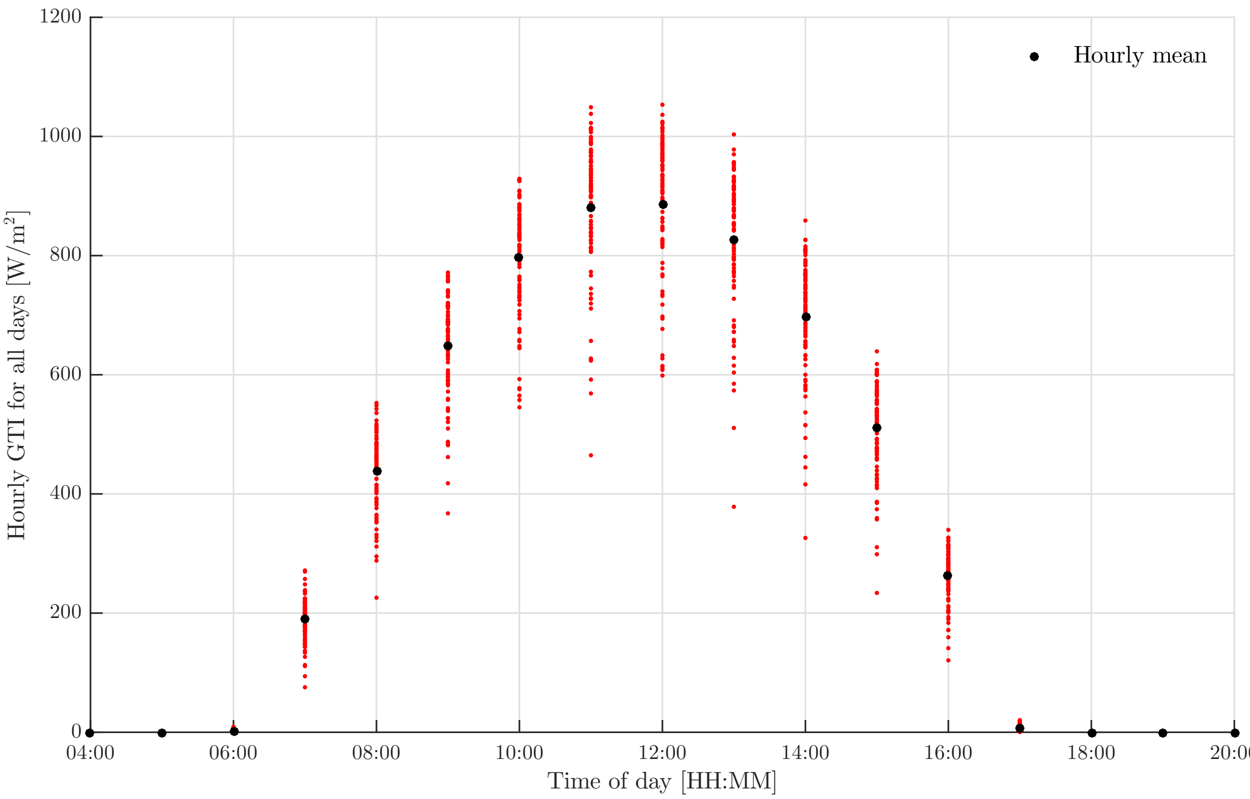

Figure 4.77: Hourly GTI values for the optimal winter distribution in site group 2 for objective function 5.

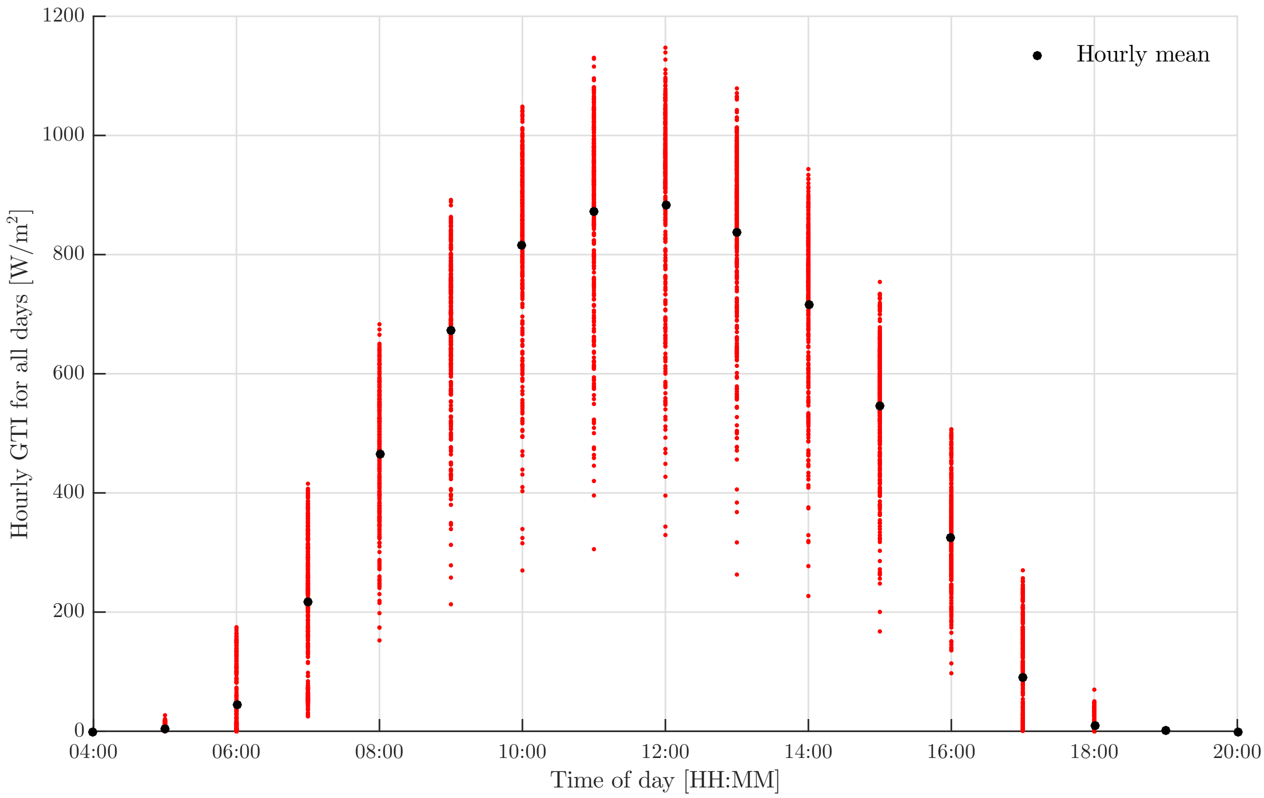

Figure 4.78: Hourly GTI values for the optimal full year distribution in site group 2 for objective function 5.

Figures 4.79 and 4.80 and Figures 4.81 and 4.82 represent the relevant probability distributions for the site group 3 winter and full year results, respectively, while Figures 4.83 and 4.84 present the corresponding hourly GTI graphs. These results show that the cumulative daily energy for the optimal winter distribution is limited within 15 % and -30 % of the mean value and goes no lower than 60 % of the maximum value throughout the season, while the cumulative daily energy for the optimal full year distribution varies between 20 % and -55 % relative to the mean value and upwards of 30 % of the maximum value throughout the year.

For the winter case, the probability distribution of cumulative daily energy relative to the maximum value shows a distinct increase in negative skewness compared to the equivalent result for objective function 4, without any increase in the maximum deviation relative to the mean or maximum values. This is reflected in the hourly GTI profile, which exhibits a slightly lower level of vertical dispersion than the equivalent profile for objective function 4. The results for the full year case also show increased negative skewness for the probability distribution relative to the maximum, though there is a corresponding increase in the maximum deviation of cumulative daily energy relative to the mean and maximum values. The hourly GTI profile shows decreased vertical dispersion (i.e. a smaller value range) for the majority of values, resulting in a more frequent occurence of outliers.

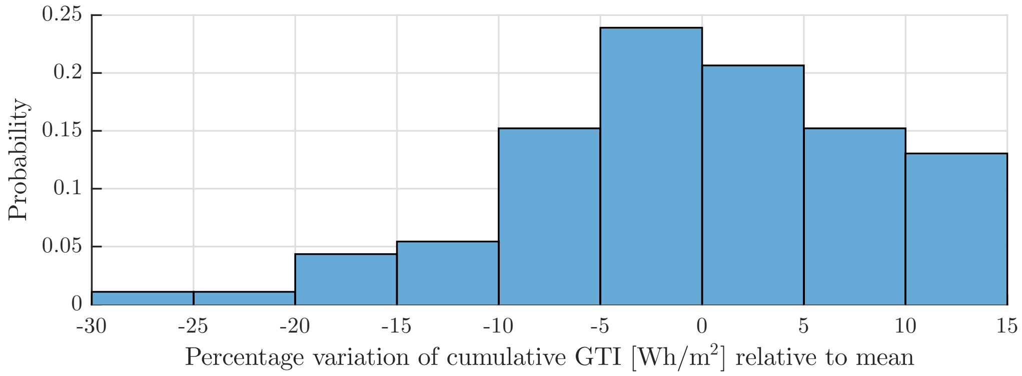

Figure 4.79: Probability distribution of cumulative daily energy relative to the mean value for the optimal winter distribution in site group 3 for objective function 5.

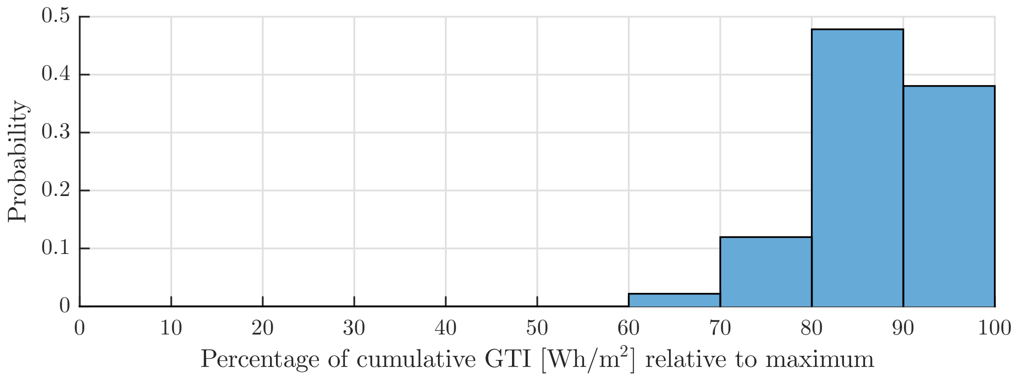

Figure 4.80: Probability distribution of cumulative daily energy relative to the maximum value for the optimal winter distribution in site group 3 for objective function 5.

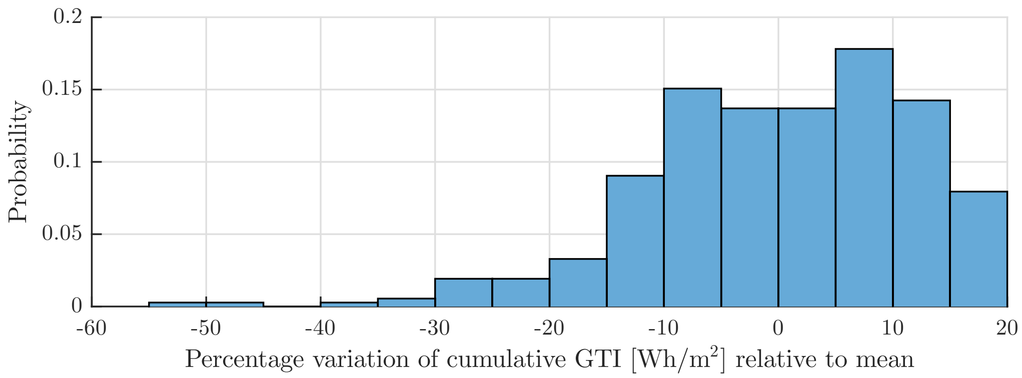

Figure 4.81: Probability distribution of cumulative daily energy relative to the mean value for the optimal full year distribution in site group 3 for objective function 5.

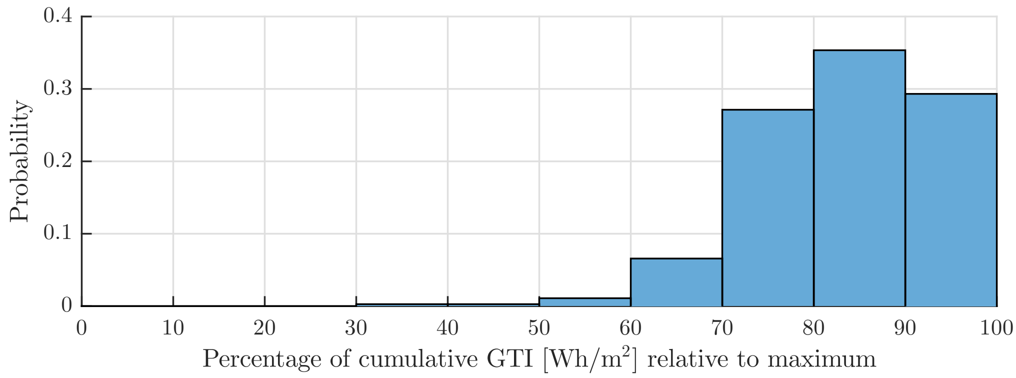

Figure 4.82: Probability distribution of cumulative daily energy relative to the maximum value for the optimal full year distribution in site group 3 for objective function 5.

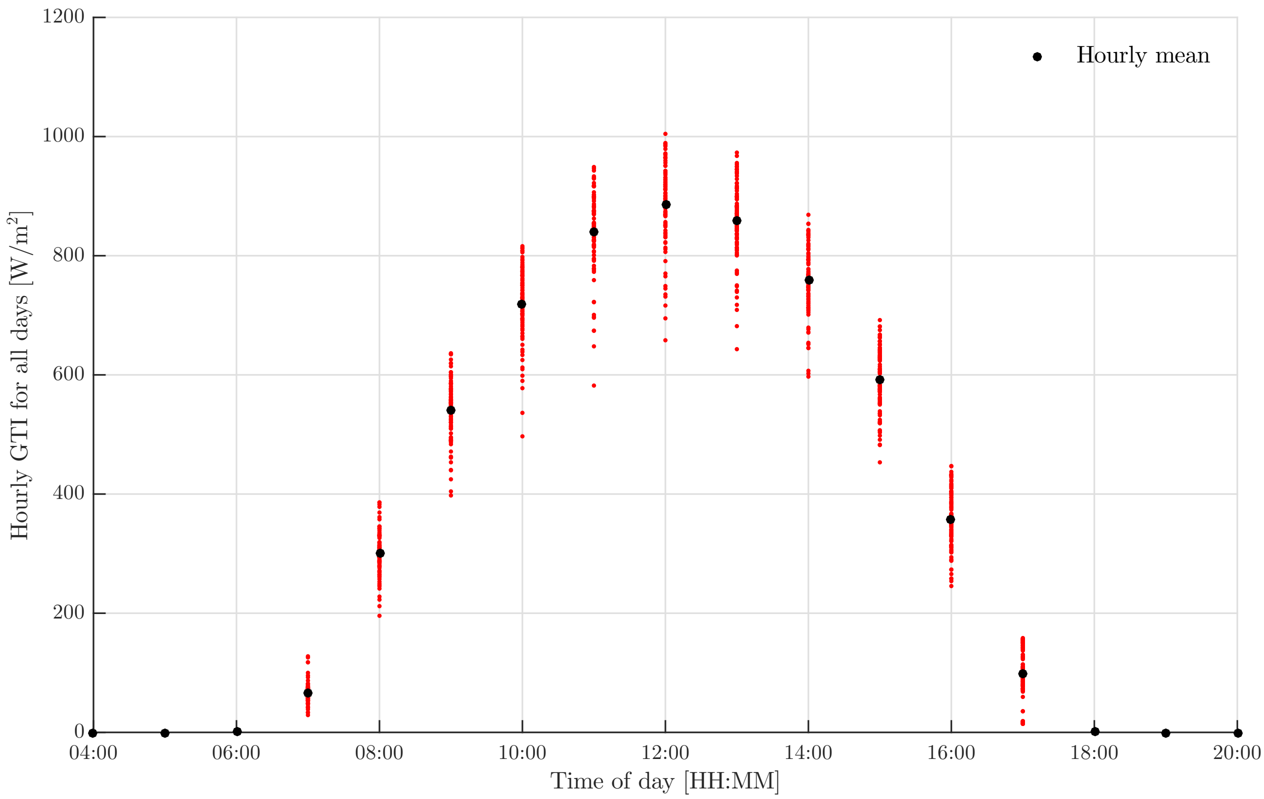

Figure 4.83: Hourly GTI values for the optimal winter distribution in site group 3 for objective function 5.

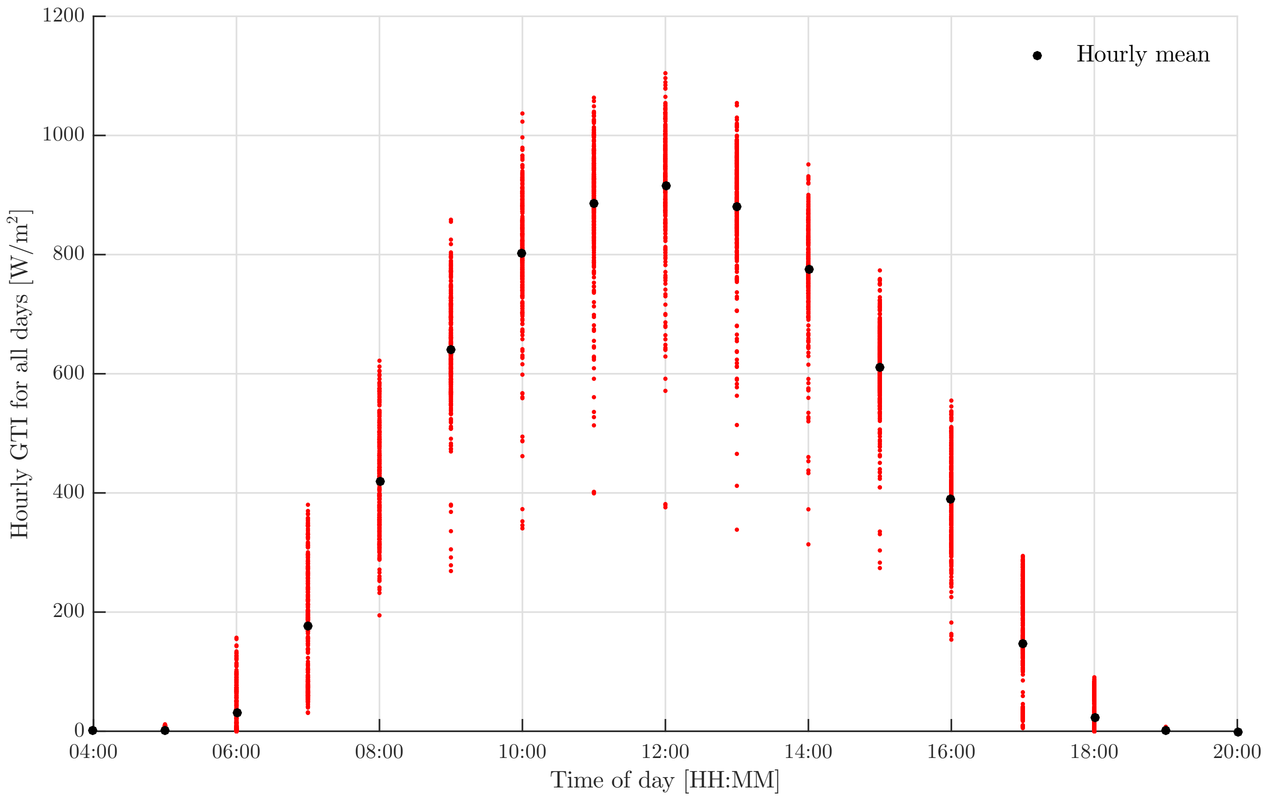

Figure 4.84: Hourly GTI values for the optimal full year distribution in site group 3 for objective function 5.

Overall, the results for problem cases optimised with this objective function confirm that it is successful in intensifying the negative skewness of the probability distribution for the cumulative daily energy relative to its maximum value. For some cases the objective function was shown to also increase the worst-case scenario, i.e. the lowest cumulative daily energy value relative to the maximum and mean values, when compared to the results for objective function 4 where the SD is minimised. This can be problematic as it pertains to the grid and its energy demand profile, and would likely necessitate the installation of a higher surplus generation capacity. However, it appears that the worst-case scenario is mitigated by the larger selection of potential locations available in site group 3, which again indicates that geographical diversity is a highly advantageous factor for controlling variability.

Chapter 5

Conclusions and recommendations

5.1 Overview

Historically, the generation of electrical energy in South Africa has relied predominately on the use of fossil fuels, particularly coal. A strong renewable energy program, mostly consisting of wind energy and photovoltaic technologies, have been implemented successfully in recent years. Due to the solar potential of the region and decreasing cost of solar PV systems, solar photovoltaic sources represent a major component of this RE fleet. To date, however, the feed-in tariffs negotiated for renewable energy sources are flat in both the seasonal and diurnal contexts. This is not compatible with the very strong pricing signals associated with consumer tariff systems such as Eskom's Megaflex tariff structure, which applies to various prominent load sectors, including the industrial sector and municipal resellers. This pricing structure is currently also being aggressively expanded to the commercial and residential sectors through the use of smart metering systems. It therefore follows that the feed-in tariffs applied to RE generation should also reflect the Megaflex tariff structure in the sense that it represents the cost of electricity generation along seasonal and diurnal timelines.

In the face of the increasing penetration of renewable energy, modern smart grid approaches to energy balance require that the mix of renewable energy (RE) technologies, e.g. wind and solar PV, should be optimised for maximum grid support. The geographical distributions of these RE plants should therefore take cognizance of the seasonal and diurnal cycles associated with both the grid load profile and the RE generation profiles. This project investigates one of these aspects, namely the long-term effects of the geographical distribution of solar PV plants on grid-related metrics such as the availability and variability of available energy in the context of relevant seasonal and diurnal considerations. The investigation was implemented as a optimisation study in the South African context, with the methodology characterised by the following aspects:

- Selection of potential photovoltaic (PV) plant locations that provide a meaningful geographical basis for the evaluation of seasonal and diurnal trends.

- Identification of a range of meaningful optimisation objectives for evaluating the significance of various grid-support metrics for long-term solar PV power generation.
- Identification of various optimisation algorithms for finding optimal distributions of PV generation capacity using objective functions derived from the aforementioned optimisation objectives.
- Design and subsequent evaluation of set of optimisation problem cases incorporating the aforementioned features.

The evaluation of the designated optimisation problem cases was implemented within an integrated software platform that required the development of a novel solar PV optimisation module. The module was integrated with the established main application interface and relational database structure of an ongoing diversified software project, and the necessary simulation software was incorporated to perform optimisation simulations on a suitable external simulation platform. The main research emphasis regarding the analysis and interpretation of the optimisation results is on the following:

- Determining the long-term seasonal and diurnal trends with regard to the aggregated solar PV power profiles and cumulative available energy associated with the distribution of generation capacity along the geographical north-south and east-west axes.
- Evaluating the potential merits of optimising the geographical allocation of solar PV generation capacity using grid-support strategies that incorporate diurnal and seasonal considerations.
- Evaluating the performance of the various optimisation algorithms in the context of the difficult multi-dimensional search spaces associated with the selected optimisation problem cases.

5.2 Conclusions

5.2.1 Integrated software platform

As prescribed by the project description, a solar PV optimisation software module was designed and utilised for the implementation of the designated optimisation study. The module was successfully consolidated with the established main application interface and relational database infrastructure of an ongoing software project, and external simulation software was incorporated for performing optimisation simulations. The relational database structure originally associated with the main application interface was evaluated and adapted where needed to provide efficient long-term storage for all data utilised and produced throughout the optimisation process. Due to its communal use in the context of the ongoing software project, the database structure was designed from a generic perspective in order to accommodate data associated with a diverse range of analytical applications.

The organisational storage capabilities offered by the relational database proved highly useful in the context of the optimisation study, both in terms of evaluating optimisation problem cases and facilitating the subsequent analysis of the optimisation results. In this regard, the graphical user interface (GUI) functionality of the solar PV optimisation module along with its built-in analytical features also proved very advantageous for the evaluation of individual problem cases. Meanwhile, the external simulation platform and its corresponding optimisation software performed well in its role of implementing the required optimisation simulations.

Although this integrated software platform was developed and used for the implementation of the described optimisation study, it can be used for the evaluation of optimisation problem cases with any seasonal characteristic, diurnal specification (where applicable), group of locations or power profile resolution. Additionally, the *type* of power profiles used, whether solar or electrical, are also not important as long as the data is of an averaged nature. This effectively means that despite its designation as a solar PV optimisation module, the implemented module can theoretically be used to optimise the distribution of generation capacity for any type of energy source. With regard to future development, the nature of the software infrastructure is such that the functionality of the solar PV optimisation module can easily be extended to include additional objective functions or optimisation algorithms for which the simulation software has been independently implemented.

5.2.2 Optimisation algorithms

For most of the problem cases evaluated in the optimisation study, the relative discrepancies between both the minima found by different optimisation algorithms, as well as the minima for different evaluations by a non-deterministic algorithm, proved to be negligible to fairly small. This is consistent with the metaheuristic characteristic of producing solutions of "good enough" quality, which approximate the global optimum without guaranteeing convergence to the true global minimum within the problem search space. It also indicates that the selected optimisation algorithms are fairly robust with regard to the diverse range of multi-dimensional search problems produced by the implementation of the different optimisation objectives.

The greatest variation in solution quality occurred for problem cases with a particular focus on the evening peak period, which can likely be ascribed to the low power levels associated with that diurnal specification. In other words, a difference that is relatively small when considered for minima based on large power or cumulative energy values becomes much more significant when a disparity of the same value occurs for minima based on low power or energy values. It can therefore be concluded that the solution quality of the optimisation algorithms deteriorates when the power levels associated with specified diurnal period are very low.

The genetic algorithm (GA) proved to be the most expensive optimisation function in terms of simulation runtime and memory requirements. For a single evaluation the cost is not too significant, but the non-deterministic nature of the technique necessitates multiple evaluations to ensure that the best solution available to the algorithm is found, resulting in a substantial cumulative duration. Due to its wider range of potential locations, which increases problem difficulty by introducing additional dimensions to the search space, the problem cases evaluated for site group 3 were also considerably more expensive than those evaluated for site groups 1 and 2. Consequently, even though it yielded high-quality solutions when a sufficient number of evaluations were considered, the GA was deemed non-ideal for the purposes of this optimisation study due to its relatively expensive time requirement and low tolerance for problem scalability.

Pattern search variations 1 and 2 performed at a similar level for the various optimisation objectives, with pattern variation 2 exhibiting slightly better solution-quality overall. This is consistent with the tendency of pattern search algorithms incorporating the generating set search (GSS) method to perform better for bounded search problems than pattern search algorithms that only use the generalised pattern search (GPS) technique. Both variations generally exhibit lower overall solution quality than the best solutions produced by both the GA and pattern variation 3, with the notable exception of the problem cases evaluated with objective function 4. In contrast to the other objective functions, for which the consistency of solution quality varies, the solutions produced for objective function 4 are characterised by very high quality and consistency. This implies that the pattern search technique is very well suited for solving search problems of the type described by objective function 4. Despite the more inconsistent solution quality exhibited for the other problem cases, however, the deterministic nature and negligible simulation runtime associated with these algorithms make them viable candidates for solving problems where approximate solutions are sufficient.

Of the four optimisation algorithms investigated, pattern search variation 3 produced the best overall solution quality over the range of objectives functions while remaining considerable less expensive than the GA in terms of simulation runtime. Its solution quality can be ascribed to the combined use of both GPS and GA search functionality, while the integrated GA implementation is simple enough to render it relatively inexpensive. As such, pattern search variation 3 was identified as the most suitable candidate for general optimisation purposes in the context of this study, given that a sufficient number of evaluations are performed to ensure that the best possible solution is found. With regard to the latter consideration, pattern search variation 3 exhibits a similar level of variation to the GA for different evaluations of the same problem case, with the exception of the problem cases evaluated for objective function 4. The consistency of solution quality in this regard can likely be improved via proper parameter adjustment for the implementation of each objective function.

5.2.3 Optimisation results

5.2.3.1 Overview

Overall, the results produced throughout the optimisation study clearly confirm the significance of diurnal and seasonal considerations for optimising the allocation of solar PV generation in the context of grid support. The results associated with the individual optimisation objectives also indicate the merits of optimising solar PV allocation for the particular grid-support aspect considered by each objective. Moreover, the observed trends strongly support the implementation of variable feed-in tariff structures for RE generation as a means of optimising the geographical distribution of solar PV generation capacity according to grid-support metrics. The frequent disparity between the results of problem cases optimised for the full year scenario compared to the results for equivalent winter cases indicate that due consideration of the high and low demand seasons is especially important in this regard.

5.2.3.2 Objective function 1: Maximisation of the daily averaged energy

The results for problem cases analysed with this optimisation objective clearly show the variation introduced in the longitudinal dimension when solar PV distribution is optimised for different diurnal specifications, as well as the variation introduced in the latitudinal dimension when considering different seasonal specifications. The distributions for maximised energy in the morning peak period or evening peak period often differ from the corresponding allocation for the full day scenario, while the same solutions also vary from season to season. For all of the analysed problem cases the optimised solutions favour a single location with all or most of the available generation capacity, which is expected for the maximisation of cumulative energy. Due to this lack of geographical diversity for a given allocation, however, the variability of the power profile and cumulative daily energy exhibited over a seasonal or annual duration is largely dictated by the often significant variability associated with a single location.

5.2.3.3 Objective function 2: Minimisation of the coefficient of variation for the daily averaged power profile

For the site groups that incorporate geographical diversity along the east-west axis, the majority of the solutions found for this optimisation objective distribute most or all of the generation capacity between two locations extending over a significant portion of the available longitudinal range. Compared to the power profiles produced by maximisation of the cumulative daily energy, the resulting aggregated daily averaged power profiles manage to reduce the midday peak power level while maintaining relatively high power levels at the start and/or the end of the diurnal cycle.

Although the optimisation objective is satisfied to a certain extent, the conceptual goal of producing an aggregated averaged power profile characterised by a wide, flat diurnal cycle is not achieved, most likely due to the limited geographical range of the country in the longitudinal dimension. As such, the merit of this optimisation objective proved somewhat limited in the local context. If applied to a grid spanning a significantly larger geographical area, however, the potential benefits offered by this optimisation objective may be amplified considerably. One excellent candidate for such an evaluation is the United States mainland, which has a latitudinal span of approximately 23.5 degrees and a longitudinal span of approximately 57 degrees, compared to the South African range of 12.6 degrees and 16.4 degrees, respectively. The inter-connected European grid presents another highly suitable option.

5.2.3.4 Objective function 3: Maximisation of the daily averaged energy weighted according to the Megaflex tariff structure

For many of the problem cases analysed for this optimisation objective, the resulting solar PV distributions are very similar to the solutions found for the equivalent problem cases optimised for the full day scenario with objective function 1. The distinct exception to this trend is the solutions for the winter period for site groups 1 and 3. This disparity between the allocations of generation capacity made for the weighted and unweighted energy maximisation objectives is highly significant as it illustrates the substantial impact of the extreme tariff associated with the peak periods of the high demand season. This characteristic is a strong indicator for the potential value of incorporating variable feed-in tariffs with diural as well as seasonal specifications for solar PV power generation. The precise structuring of such tariffs should be investigated further in the context of both plant profitability and long-term grid-support that is complementary to RE generation profiles from other sources.

5.2.3.5 Objective function 4: Minimisation of the relative standard deviation of the cumulative daily energy

This optimisation function was highly successful in producing distributions with low variability of the cumulative daily energy values presented for an entire seasonal or annual duration. For most of the problem cases, the daily variability for the resulting distribution is significantly lower compared to that of corresponding distributions found for maximised energy levels, but comes at the cost of lower cumulative energy delivered over the course of a season or year. For the first time, the allocation of generation capacity is highly diversified and seasonal trends become less obvious. However, the significance of a seasonal consideration is again confirmed by the fact that the problem cases optimised for the full year scenario do not reflect the optimal solutions for the winter scenario and *vice versa*. The lowest variability is also consistently exhibited by the distributions found for site group 3, which supports the notion that an increase in the geographical range and intensity of diversification serves to reduce variability for aggregated solar PV power profiles.

5.2.3.6 Objective function 5: Maximisation of the negative skewness of the probability distribution for the cumulative daily energy

The distributions produced for this optimisation objective were successful in skewing variability throughout the relevant seasonal period so that the cumulative available energy for the majority of days is in the region of the maximum value for any one day. The distributions still show a significant level of diversification, but tends to favour locations with high cumulative annual or seasonal energy yields. As such, the variability for these distributions is consistently worse than for equivalent solutions found for objective function 4, exhibiting much more frequent and extreme outlier values on a daily and hourly basis. On the other hand, the cumulative energy yield on an annual or seasonal basis is generally notably higher than that of the distributions produced by objective function 4. The same applies to the cumulative yield for the daily averaged power profile.

5.3 Recommendations

5.3.1 Optimisation algorithms

The following research aspects regarding the use of optimisation algorithms in the context of this study are recommended for further exploration:

- Implementation of test problem cases with larger groups of locations in order to evaluate the performance of the pattern search variations for increased problem difficulty.
- Parameter adjustment of pattern variation 3 to maximise solution quality and consistency for the types of search problems associated with each of the objective functions implemented in this study.
- Comparative evaluation of the performance of single-solution metaheuristic techniques other than pattern search in the context of this optimisation study. Suitable algorithms for this purpose include simulated annealing, tabu search and variable neighbourhood search.
- Comparative evaluation of the performance of population-based metaheuristic techniques other than GA in the context of this optimisation study. Suitable algorithms for this purpose include particle swarm optimisation and scatter search.
- Implementation of promising metaheuristic techniques as combined algorithms similar to pattern search variation 3, followed by comparative evaluation in the context of this optimisation study.

5.3.2 Optimisation strategy

This project by no means offers a complete solution to the complex problems surrounding grid-integration of large-scale RE generation, but it does provide a good foundation for the development of comprehensive models for optimising the allocation of RE generation capacity. The following research is recommended with regard to further evaluation, refinement and development of the optimisation methodology investigated in this project:

- Implementation of the described optimisation strategy using the locations of all the existing and planned solar PV installations in South Africa, with subsequent evaluation of the current allocation of PV generation capacity compared to scenarios optimised for the grid-support.

- Evaluation of the described optimisation strategy in the context of a power grid spanning a larger geographical area.
- Implementation of the described optimisation strategy using high-quality measured solar data in order to validate the trends exhibited by the results produced for the synthetically generated solar power profiles.
- Comparative analysis of measured solar PV power generation and local solar radiation data to evaluate the suitability of predicting long-term solar PV performance characteristics via solar power profiles.
- Development and implementation of comprehensive models that incorporate plant parameters and additional environmental parameters for more accurate estimation of the power output profiles for hypothetical solar PV plants.
- Evaluation of alternative approaches for reducing the variability of solar PV power generation, such as the cost-variance analysis method commonly used for optimising wind farm allocation [73, 74].
- Implementation of hybridised optimisation of generation capacity allocation for a mix of solar PV and wind energy.
- Investigation of multi-objective optimisation for RE power generation that combines variability minimisation and maximisation of energy availability for high demand periods.
- Comprehensive evaluation of the effects of variable feed-in tariff structures for optimising the allocation of RE generation capacity, both in terms of grid support and plant profitability.

Bibliography

[1] A. Jager-Waldau, “Photovoltaics and renewable energies in Europe,” *Renew. Sustain. Energy Rev.*, vol. 11, no. 7, pp. 1414–1437, 2007.

[2] A. Eberhard, J. Leigland, and J. Kolker, “South Africa’s Renewable Energy IPP Procurement Program: Success Factors and Lessons,” PPIAF, Washington, Tech. Rep., 2014. [Online]. Available: http://www.ee.co.za/article/south-africas-reippp-programme-success-factors-lessons.html

[3] D. Newbery and A. Eberhard, “South African Network Infrastructure Review: Electricity,” Tech. Rep., 2008. [Online]. Available: http://www.gsb.uct.ac.za/files/saelectricitypaper08.pdf

[4] T. Matona, “Power System Status Update,” Presentation, Eskom, 2014.

[5] D. I. Stern, “Energy and economic growth in the USA,” *Energy Econ.*, vol. 15, no. 2, pp. 137–150, 1993.

[6] “State of Renewable Energy in South Africa,” Republic of South Africa: Department of Energy, Tech. Rep., 2015.

[7] “Integrated Resource Plan for Electricity 2010-2030,” 2011.

[8] D. Fourie, L. Kritzinger-van Niekerk, and M. Nel. (2015, June) An overview of the renewable energy independent power producers procurement programme (REIPPPP).

[9] “Global Trends in Renewable Energy Investmenet 2015,” Frankfurt School of Finance and Management, Tech. Rep., 2015.

[10] M. Rycroft, “Summary of REIPPP round three projects,” *Energize*, pp. 7–9, dec 2013. [Online]. Available: http://www.ee.co.za/wp-content/uploads/legacy/Energize_2013/03_iDn_summary-of-reippp.pdf

[11] (2015) Bid window 4 total contracted capacity from Preferred Bidders. Republic of South Africa: Department of Energy.

[12] IRENA, “Renewable Power Generation Costs in 2014,” International Renewable Energy Agency, Tech. Rep., 2015. [Online]. Available: www.irena.org/publications

[13] K. Branker, M. J. M. Pathak, and J. M. Pearce, "A review of solar photovoltaic levelized cost of electricity," *Renew. Sustain. Energy Rev.*, vol. 15, no. 9, pp. 4470–4482, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2011.07.104

[14] N. Mithulananthan, R. Shah, and K. Y. Lee, "Small-disturbance angle stability control with high penetration of renewable generations," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1463–1472, 2014.

[15] H. Chen, "Key technologies for renewable energy integration - A full scale demonstration at Hainan Island," in *2013 IEEE PES Asia-Pacific Power Energy Eng. Conf.*, 2013.

[16] E. M. G. Rodrigues, A. W. Bizuayehu, and J. P. S. Catalão, "Analysis of Requirements in Insular Grid Codes for Large-Scale Integration of Renewable Generation," in *2014 IEEE/PES Transm. Distrib. Conf. Expo.*, 2014.

[17] M. Milligan, K. Porter, E. DeMeo, P. Denholm, H. Holttinen, B. Kirby, N. Miller, A. Mills, M. O'Malley, M. Schuerger, and L. Söder, "Preface: Wind Power Myths Debunked," in *Wind Power Power Syst.*, 2nd ed., T. Ackermann, Ed. Chichester: John Wiley Sons. Ltd, 2012, ch. 2.

[18] P. Capros, N. Tasios, and A. Marinakis, "Very High Penetration of Renewable Energy Sources to the European Electricity System in the context of model-based analysis of an energy roadmap towards a low carbon EU economy by 2050," in *2012 9th Int. Conf. Eur. Energy Mark.*, 2012, pp. 1–8.

[19] D. A. Halamay, T. K. A. Brekken, A. Simmons, and S. McArthur, "Reserve requirement impacts of large-scale integration of wind, solar, and ocean wave power generation," *IEEE Trans. Sustain. Energy*, vol. 2, no. 3, pp. 321–328, 2011.

[20] M. Bello, C. Carter-Brown, R. Smit, and I. E. Davidson, "Power Planning for renewable energy grid integration - Case Study of South Africa," in *IEEE Power Energy Soc. Gen. Meet.*, 2013, pp. 1–5.

[21] H. Chen, S. Member, P. Xuan, Y. Wang, K. Tan, and X. Jin, "Key Technologies for Integration of Multitype Renewable Energy Sources Research on Multi-Timeframe Robust Scheduling / Dispatch," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 471–480.

[22] Z. R. Z. Ruihua, D. Y. D. Yumei, and Y. L. Y. Liu, "New challenges to power system planning and operation of smart grid development in China," in *Int. Conf. Power Syst. Technol.*, 2010, pp. 1–8.

[23] A. Datta, P. Mohanty, and M. Gujar, "Accelerated deployment of Smart Grid technologies in India - Present scenario, challenges and way forward," *2014 IEEE PES Innov. Smart Grid Technol. Conf.*, 2014.

[24] S. Sinha and S. Chandel, "Review of recent trends in optimization techniques for solar photovoltaic-wind based hybrid energy systems," *Renew. Sustain. Energy Rev.*, vol. 50,

pp. 755–769, 2015. [Online]. Available:
http://linkinghub.elsevier.com/retrieve/pii/S1364032115005006

[25] N. S. Thomaidis, F. J. Santos-Alamillos, D. Pozo-Vázquez, and J. Usaola-García, "Optimal management of wind and solar energy resources," *Comput. Oper. Res.*, vol. 66, pp. 284–291, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0305054815000556

[26] G. R. G. Hoste, M. J. Dvorak, and M. Z. Jacobson, "Matching Hourly and Peak Demand by Combining Different Renewable Energy Sources: A Case Study for California in 2020," Stanford University: Department of Civil and Environmental Engineering, Tech. Rep., 2009.

[27] B. K. Sovacool, "The intermittency of wind, solar, and renewable electricity generators: Technical barrier or rhetorical excuse?" *Util. Policy*, vol. 17, no. 3-4, pp. 288–296, 2009.

[28] P. J. D. Lloyd, "The Renewable Energy Adventure and a Transition to a Low-Carbon South Africa," in *2012 Proc. 9th Ind. Commer. Use Energy Conf.*, 2012, pp. 1–8.

[29] M. Gomez, F. Jurado, P. Diaz, and N. Ruiz-Reyes, "Evaluation of a Particle Swarm Optimization Based Method for Optimal Location of Photovoltaic Grid-connected Systems," *Electr. Power Components Syst.*, vol. 38, no. 10, pp. 1123–1138, 2010.

[30] P. D. Huy, V. K. Ramachandaramurthy, and M. P. H. A, "Optimal Selection of Location, Sizing and Power factor for Solar PV Plants using Differential Evolution," in *IEEE Innov. Smart Grid Technol. - Asia*, 2015.

[31] B. Urquhart, M. Sengupta, and J. Keller, "Optimizing geographic allotment of photovoltaic capacity in a distributed generation setting," *Prog. Photovoltaics Res. Appl.*, vol. 21, no. 6, pp. 1276–1285, 2013.

[32] J. Arán Carrión, a. Espín Estrella, F. Aznar Dols, M. Zamorano Toro, M. Rodríguez, and a. Ramos Ridao, "Environmental decision-support systems for evaluating the carrying capacity of land areas: Optimal site selection for grid-connected photovoltaic power plants," *Renew. Sustain. Energy Rev.*, vol. 12, no. 9, pp. 2358–2380, 2008.

[33] M. Pöller, M. Ober, and G. Moodley, "Analysis of options for the future allocation of PV farms in South Africa," Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Tech. Rep., 2015.

[34] "Variability of Wind Power and other Renewables: Management Options and Strategies," International Energy Agency, Tech. Rep., 2005. [Online]. Available: http://www.iea.org/papers/2005/variability.pdf

[35] B. Parsons, M. Hummon, J. Cochran, B. Stoltenberg, P. Batra, B. Mehta, and D. Patel, "Variability of Power from Large- Scale Solar Photovoltaic Scenarios in the State of Gujarat," National Renewable Energy Laboratory (NREL), Tech. Rep., 2014.

[36] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Prog. Energy Combust. Sci.*, vol. 34, no. 5, pp. 574–632, 2008.

[37] "Tariffs Charges Booklet 2014/15," Eskom, 2014.

[38] "Eskom's Tariff Design Methodology 2008/2009," Eskom, Tech. Rep., 2008.

[39] G. M. Masters, *Renewable and Efficient Electric Power Systems.* Hoboken: Wiley-Interscience, 2004.

[40] Sun path chart program. University of Oregon: Solar Radiation Monitoring Laboratory. [Online]. Available: http://solardat.uoregon.edu/SunChartProgram.html

[41] C. Xiao, X. Yu, D. Yang, and D. Que, "Impact of solar irradiance intensity and temperature on the performance of compensated crystalline silicon solar cells," *Solar Energy Materials Solar Cells*, vol. 128, pp. 427–434, 2014.

[42] D. R. Myers, *Solar Radiation: Practical Modeling for Renewable Energy Applications*, A. Ghassemi, Ed. Boca Raton: CRC Press, 2013.

[43] M. Gulin, M. Vasak, and M. Baotic, "Estimation of the global solar irradiance on tilted surfaces," in *17th Int. Conf. Electr. Drives Power Electron. (EDPE 2013)*, 2013, pp. 334–339.

[44] (2015) SolarGIS homepage. SolarGIS. [Online]. Available: http://solargis.com/products/maps-and-gis-data/free/download/

[45] E. Drury, A. Lopez, P. Denholm, and R. Margolis, "Relative performance of tracking versus fixed tilt photovoltaic systems in the USA," *Prog. Photovolt Res. Appl.*, vol. 22, pp. 1302–1315, 2013. [Online]. Available: http://dx.doi.org/10.1002/pip.1160

[46] J. Isotalo, *Basics of Statistics.* Tampere: CreateSpace Independent Publishing Platform.

[47] A. Messac, *Optimization in Practice with MATLAB.* New York: Cambridge University Press, 2015.

[48] F. Rothlauf, *Design of Modern Heuristics: Principles and Applications*, G. Rozenberg, T. Back, J. Kok, and H. Spaink, Eds. Heidelberg: Springer, 2011.

[49] P. Siarry and Z. Michalewicz, *Advances in metaheuristics for hard optimization.* New York: Springer, 2008.

[50] F. Glover and G. A. Kochenberger, *Handbook of metaheuristics.* Boston: Kluwer Academic Publishers, 2003.

[51] A. Fraser, "Simulation of Genetic Systems by Automatic Digital Computers," *Aust. J. Biol. Sci.*, vol. 11, no. 4, pp. 603–612, 1958.

[52] H. J. Bremermann, "Optimization through evolution and recombination," in *Proc. Conf. Self-Organizing Syst. – 1962*, M. Yovits, G. Jacobi, and G. Goldstine, Eds., 1962, pp. 93–106.

[53] J. Holland, *Adaptation in Natural and Artificial Systems.* Cambridge: MIT Press, 1975.

[54] K. Man, T. Tang, and S. Kwong, *Genetic algorithms: concepts and designs.* London: Springer, 1999.

[55] M. Gen, R. Cheng, and L. Lin, *Network models and optimization: multiobjective genetic algorithm approach.* London: Springer, 2008.

[56] E. K. Burke and G. Kendall, *Search Methodol. Introd. Tutorials Optim. Decis. Support Tech.* New York: Springer, 2005, ch. 4, pp. 97–126.

[57] V. Torczon, "On the Convergence of Pattern Search Algorithms," *SIAM J. Optim.*, vol. 7, no. 1, pp. 1–25, 1997.

[58] R. M. Lewis and V. Torczon, "Pattern Search Algorithms for Bound Constrained Minimization," *SIAM J. Optim.*, vol. 9, no. 4, pp. 1082–1099, 1999.

[59] ——, "Pattern Search Methods for Linearly Constrained Minimization," *SIAM J. Optim.*, vol. 10, no. 3, pp. 917–941, 2000.

[60] C. Audet and J. Dennis Jr., "Analysis of Generalized Pattern Searches," *SIAM J. Optim.*, vol. 13, no. 3, pp. 889–903, 2003.

[61] R. M. Lewis, A. Shepherd, and V. Torczon, "Implementing Generating Set Search Methods for Linearly Constrained Minimization," *SIAM J. Sci. Comput.*, vol. 29, no. 6, pp. 2507–2530, 2007.

[62] Oracle, "A Relational Database Overview," 2015. [Online]. Available: https://docs.oracle.com/javase/tutorial/jdbc/overview/database.html

[63] G. Powell, *Beginning Database Design.* Indianapolis: Wiley Publishing, Inc., 2006.

[64] P. Beynon-Davies, *Database Systems*, 3rd ed. New York: Palgrave Macmillan, 2004.

[65] E. F. Codd, "A relational model of data for large shared data banks," *Commun. ACM*, vol. 26, no. 1, pp. 64–69, 1970.

[66] (2015) The 10 Most Popular DB Engines (SQL and NoSQL) in 2015. [Online]. Available: https://blog.jooq.org/2015/10/15/the-10-most-popular-db-engines-sql-and-nosql-in-2015

[67] J. Mack. (2014) Five Advantages Disadvantages Of MySQL. [Online]. Available: https://www.datarealm.com/blog/five-advantages-disadvantages-of-mysql/

[68] B. Bekker, "Irradiation and PV array energy output, cost and optimal positioning estimation for South Africa," *J. Energy South. Africa*, vol. 18, no. 2, pp. 16–25, 2007.

[69] R. Aguiar, M. Collares-Pereira, and J. Conde, "Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices," *Sol. Energy*, vol. 40, no. 3, pp. 269–279, 1988.

[70] R. Aguiar and M. Collares-Pereira, "A Time-Dependent, Autoregressive, Gaussian Model for Generating Synthetic Hourly Radiation," *Sol. Energy*, vol. 49, no. 3, pp. 167–174, 1992.

[71] *PVsyst User's Manual*, PVsyst SA, Satigny, 2014.

[72] *Global Optimization Toolbox: User's Guide*, MathWorks, 2016.

[73] Y. Degeilh and C. Singh, "A quantitative approach to wind farm diversification and reliability," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 2, pp. 303–314, 2011.

[74] B. Drake and K. Hubacek, "What to expect from a greater geographic dispersion of wind farms?-A risk portfolio approach," *Energy Policy*, vol. 35, no. 8, pp. 3999–4008, 2007.